Analysis and Design of Vehicular Networks

A Thesis
Presented to
The Academic Faculty

by

``

Hao Wu

In Partial Fulfillment
Of the Requirements for the Degree of
Doctor of Philosophy in Computer Science

Georgia Institute of Technology
December, 2005

Analysis and Design of Vehicular Networks

Approved By:

Professor Richard Fujimoto
Advisor, College of Computing
Georgia Institute of Technology

Professor Mostafa Ammar
College of Computing
Georgia Institute of Technology

Professor Randall Guensler
College of Engineering
Georgia Institute of Technology

Professor Michael Hunter
College of Engineering
Georgia Institute of Technology

Professor George Riley
College of Engineering
Georgia Institute of Technology

Date Approved 9/14/2005

To my wife and my parents,

For their love and unreserved support

They are the source of energy that drives me forward

ACKNOWLEDGEMENTS

I owe a lot of thanks to many people for the completion of this thesis. This thesis is the result of interdisciplinary collaboration. Without their help and support, I would not be able to finish this thesis work.

First, I feel greatly grateful to my advisor – Dr. Richard Fujimoto. Dr. Fujimoto led me all the way during my Ph.D. study for a period of 6 years. His advisories spanned a wide range from improving my English to basic problem-solving skills in conducting research, not to mention the financial support he provided during these years. He has been very supportive of the research topics that I chose to pursue. I feel very lucky to have Dr. Fujimoto as my advisor.

Many faculties at Georgia Tech have given me significant directions. Dr. Randall Guensler and Dr. Michael Hunter from School of Civil and Environmental Engineering have worked side-by-side with me for most works in this thesis. Dr. Mostafa Ammar from College of Computing and Dr. George Riley from School of Electrical and Computer Engineering have given me a lot of suggestions concerning networking topics.

Several graduate students have contributed directly to some works in this thesis. They are Mahesh Palekar, Jaesup Lee and Joonho Ko.

I also benefit a lot from discussion with a lot of people in College of Computing. They are Steve Ferenci, Kalyan Perumalla, Zhongtang Cai, Wenrui Zhao, Qi He, Meng Guo, just to mention a few.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS……………………………………………………………..iv

LIST OF TABLES…………………………………………………………….... ix

SUMMARY… …………………………………………………………………... xii

1 INTRODUCTION 1

1.1 Vehicular Networks 1

1.2 Infrastructure-less V2V Networks 5

1.3 Infrastructure-based Vehicular Networks 7

1.4 System Evaluation Methodology 9

1.5 Related Work 13

1.5.1 Wireless Networking 13

1.5.2 Vehicular Networks 18

1.5.3 Vehicle Traffic Flow Theory 19

1.6 Research Contributions 21

1.7 Roadmap for This Thesis 22

2 SPATIAL PROPAGATION OF INFORMATION 24

2.1 Introduction 24

2.2 Analytical Models 25

2.2.1 Definitions and Assumptions 25

2.2.2 Qualitative Analysis 29

2.2.3 Notation 32

2.2.4 Sparse Network Model 33

2.2.5 Dense Network Model 36

2.2.6 Between "Sparse" and "Dense" 41

2.2.7 Numerical Results 42

2.3 Models for Two-way Vehicle Traffic 45

2.3.1 Two-way Vehicle Traffic 45

2.3.2 Sparse Network Model 46

2.3.3 Dense Network Model 46

2.3.4 Verification 49

2.3.5 Discussion 50

2.4 Simulation Study 51

2.4.1 Experimental Design 51

2.4.2 Results 52

2.4.3 Discussion of Results 56

2.4.4 Lane Closing Message Propagation 57

2.5 Conclusion 59

3 DATA DISSEMINATION ALGORITHMS 61

3.1 Motivation 61

3.2 Data Dissemination Services 62

3.3 Enhanced Opportunistic Forwarding Design 64

3.3.1 Assumptions 65

3.3.2 A Generic Design Methodology 65

3.4 MDDV 69

3.4.1 Vehicle Traffic Information 69

3.4.2 Message Head Pair 70

3.4.3 Data Exchange 73

3.5 Optimistic Forwarding 75

3.6 Propagation Phase 78

3.7 Evaluation 78

3.7.1 Simulation Methodology 79

3.7.2 Baseline Scenario 80

3.7.3 Performance Comparison 82

3.7.4 Message Overhead 85

3.8 Discussion 86

3.8.1 Design Space 86

3.8.2 Extension 88

3.9 Summary 89

4 INFRASTRUCTURE-BASED VEHICULAR NETWORKS 90

4.1 Introduction 90

4.2 Evaluation 94

4.2.1 Test Scenario 94

4.2.2 Analytical Models 97

4.2.3 Simulation Settings 101

4.2.4 Required WWAN Capacity 103

4.2.5 WLAN Coverage and Average Data Traffic Load 105

4.2.6 Usable WLAN Capacity 108

4.2.7 Impact of Vehicle Traffic Dynamics 112

4.2.8 Number of WLAN Access Points 116
4.2.9 Data Throughput Per Vehicle 118
4.3 Implication 120
4.4 Conclusion 123
5 MEASURING SHORT RANGE COMMUNICATIONS 125
5.1 Introduction 125
5.2 Experimental Design 125
5.3 Vehicle-to-Roadside Communications 128
5.3.1 Without vs. With External Antenna 129
5.3.2 A Detailed Scenario 130
5.3.3 Roadside Station Configurations 132
5.3.4 Discussion 135
5.4 Vehicle-to-Vehicle Communications 136
5.4.1 Car Following Scenario 136
5.4.2 Car Crossing Scenario 138
5.5 Multi-hop Communication Experiments 139
5.6 Discussion 142
5.7 Conclusion 144
6 CONCLUSION 145
6.1 Vehicular Network Evaluation 145
6.2 Design of Effective Data Services 148
6.3 Future Direction 149
REFERENCES 153

LIST OF FIGURES

Figure 1: A notational vehicular network 4

Figure 2: Vehicle traffic model 12

Figure 3: An example of message propagation 26

Figure 4: Catch-up process 30

Figure 5: “Re-overtake” scenario 33

Figure 6: Average message head speed 36

Figure 7: Average partition tail speed 39

Figure 8: Numerical results 43

Figure 9: Beginning of a catch-up process 46

Figure 10: Simulation verification 50

Figure 11: I-75 corridor study area 51

Figure 12: Evening peak end-to-end propagation delay 52

Figure 13: Nighttime end-to-end propagation delay 52

Figure 14: End-to-end delay distribution 54

Figure 15: Evening peak end-to-end connectivity 55

Figure 16: Evening peak end-to-end connectivity for fleet penetration ratio of 0.1 55

Figure 17: Cumulative distribution of end-to-end delays 58

Figure 18: Geographical-temporal multicast 64

Figure 19: Lack of perfect knowledge 64

Figure 20: Contention-based scheme 76

Figure 21: Cumulative distribution of dissemination delays 81

Figure 22: Performance comparison 83
Figure 23: Normalized message overhead 85
Figure 24: One-dimensional forwarding model 98
Figure 25: Average WWAN data traffic load 104
Figure 26: Normalized WLAN coverage 106
Figure 27: The average data traffic load of a WLAN access point 108
Figure 28: Illustration of multi-hop interference 109
Figure 29: The average WLAN access point throughput 110
Figure 30: The loss ratio of the WLAN data traffic 111
Figure 31: Normalized standard deviation 113
Figure 32: Normalized Allan deviation 115
Figure 33: Number of access points required 117
Figure 34: The average vehicle throughput 120
Figure 35: Road illustration 127
Figure 36: The communication performance w/ and w/o an antenna on the vehicle 129
Figure 37: The V2R communication performance 130
Figure 38: Illustration of road geometry 130
Figure 39: Lap-by-lap results 132
Figure 40: The impact of the placement of the roadside station 133
Figure 41: Side view of the road (antenna on the table) 133
Figure 42: The impact of the height of the antenna 134
Figure 43: Side view of the road (antenna on the tripod) 134
Figure 44: The impact of the placement of the roadside station 135

Figure 45: The V2V communication performance when vehicles were moving south . 136

Figure 46: The V2V communication performance when vehicles were moving north.. 137

Figure 47: The cumulative distribution of effective communication ranges.................. 139

Figure 48: The comparison of one-hop and two-hop communication............................ 141

SUMMARY

Advances in computing and wireless communication technologies have increased interest in “*smart*” vehicles – vehicles equipped with significant computing, communication and sensing capabilities to provide services to travelers. Smart vehicles can be exploited to improve driving safety and comfort as well as optimize surface transportation systems. Wireless communications among vehicles and between vehicles and roadside infrastructure represent an important class of vehicle communications.

One can envision creating an integrated radio network leveraging various wireless technologies that work together seamlessly. Based on cost-performance tradeoffs, different network configurations may be appropriate for different environments. In urban areas, for example, it is reasonable to assume that wireless infrastructures are deployed to provide ubiquitous connectivity in a cost-effective manner. In rural areas, it might be more economical to rely on infrastructure-less vehicle-to-vehicle communications supplemented by limited infrastructures to cover hot spots or other areas of particular interest. An understanding of the properties of different vehicular network architectures, e.g., delay, loss rate, throughput, traffic load, etc., is necessary before services can be successfully deployed. Based on this understanding, efficient data services (e.g., data dissemination services) can be designed to accommodate application requirements.

This thesis examines several research topics concerning both the evaluation and design of vehicular networks.

We explore the properties of vehicle-to-vehicle (V2V) communications. V2V communications are an important component of many vehicular networks. However, V2V

communications have certain limitations, e.g., unreliable channels, dynamic topology, etc. We study the spatial propagation of information along the road using V2V communications. Our analysis identifies the vehicle traffic characteristics that significantly affect information propagation. We also evaluate the feasibility of propagating information along a highway.

There are many design alternatives existing to build infrastructure-based vehicular networks, several of which have been evaluated in a realistic vehicular environment. Based on these evaluations, we have developed some insights into the design of future broadband vehicular networks capable of adapting to varying vehicle traffic conditions.

Based on the above analysis, opportunistic forwarding that exploits vehicle mobility to overcome vehicular network partitioning appears to be a viable approach for data dissemination using V2V communications for applications that can tolerate some data loss and delay. We introduce a methodology to design enhanced opportunistic forwarding algorithms. Practical algorithms derived from this methodology have exhibited different performance/overhead tradeoffs.

An in-depth understanding of wireless communication performance in a vehicular environment is necessary to provide the groundwork for realizing reliable mobile communication services. We have conducted an extensive set of field experiments to examine the performance of short-range communications between vehicles and between vehicles and roadside stations in a specific highway scenario.

CHAPTER

1 INTRODUCTION

This chapter begins with an overview of vehicular networks. Two major types of vehicular networks – infrastructure-less and infrastructure-based, are described. We then summarize the evaluation methodology used throughout this research. Related work is introduced next. Finally we summarize the research contributions of this thesis.

1.1 Vehicular Networks

Advances in computing and wireless communication technologies have increased interest in “*smart*” vehicles – vehicles equipped with significant computing, communication and sensing capabilities to provide services to travelers. Smart vehicles are often associated with the emergence of Intelligent Transportation System (ITS) [1] with the goal of improving safety, and reduce congestion and pollution. Smart vehicles create the opportunity to deploy a plethora of new services. Applications using these in-vehicle systems can be generally classified as either safety or non-safety related. Safety applications [59, 93] include collision avoidance, cooperative driving, etc. Non-safety applications include traveler information support [99] [85], toll service, Internet access [43], or entertainment, to mention a few. The USDOT’s ITS Vehicle-Infrastructure Integration (VII) initiative is attempting to capitalize on “smart” vehicles by encouraging public-private partnerships where wireless communication devices are installed in the nation’s vehicle fleet (private investment) and roadside communication infrastructure is installed along the highways, arterials, and intersections of the transportation system (public investment) [83, 84]. Such a system has the potential to vastly improve safety [59, 93], vehicle mobility, and provide new public and commercial services[43, 85, 99]. These opportunities are near at hand, with potentially 10% of the nation's vehicle fleet instrumented within two years of the commitment to deploy the system [84].

In-vehicle systems offer the potential to greatly lower system operating costs and lessen dependence on government-maintained infrastructures. They allow coverage to extend beyond the extent of roadside infrastructure, e.g., roadside cameras that are expensive to deploy and maintain. Subject to privacy considerations, in-vehicle sensors offer the potential for much more detailed, accurate information (e.g., on-road vehicle activity and emissions) than would otherwise be impossible, enabling new ways to improve and optimize the transportation system as well as support a variety of commercial applications. In-vehicle computing systems facilitate the customization of information services to the needs and characteristics of individual travelers.

Wireless communication is clearly one of the key enabling technologies for the afore-mentioned applications. Dedicated Short Range Communications (DSRC) has been proposed to support public safety and private operations for *vehicle-to-vehicle (V2V)* and *vehicle-to-roadside (V2R)* communications. DSRC together with cellular communications provide network connectivity to moving vehicles. There are currently multiple DSRC standards programs in progress worldwide. In North America, the Federal Communication Commission (FCC) has allocated 75MHz of spectrum at 5.9GHz for DSRC [6] based on IEEE 802.11a.

A *vehicular network* is a communication network organizing and connecting "smart" vehicles to each other and mobile and fixed-location resources. A vehicular network consists of instrumented vehicles (and) network infrastructure. At a minimum, an instrumented vehicle is equipped with on-board computing, wireless communication devices, and a GPS device enabling the vehicle to track its spatial and temporal trajectory. Vehicle instrumentation may also include a pre-stored digital map and sensors for reporting crashes, engine operating parameters, etc. One cannot assume that every vehicle will have these capabilities; due to the gradual nature of market penetration, only a fraction of the vehicles on the road will be instrumented, at least for the next several years. Specifically, the term *penetration ratio* is defined as the fraction of vehicles on the road that are instrumented. In the remainder of this thesis, if not specified otherwise the term "vehicle" refers to instrumented vehicles only. Vehicular networks differ from other wireless networks (e.g., sensor networks) primarily due to the fact that users reside inside vehicles.

Several wireless technologies exist for creating vehicular networks, including Wireless Wide Area Networks (WWAN), Wireless Metro Area Networks (WMAN), Wireless Local Area Networks (WLAN) using roadside access points, and ad hoc networks using V2V communications. These technologies offer different tradeoffs in cost and performance. By utilizing these technologies, there are many ways to construct vehicular networks. Three alternatives include a pure wireless V2V ad-hoc network, a wired backbone with wireless last-hop, or a hybrid architecture using V2V communication that does not rely on a fixed infrastructure, but can exploit it for improved performance and functionality when it is available.

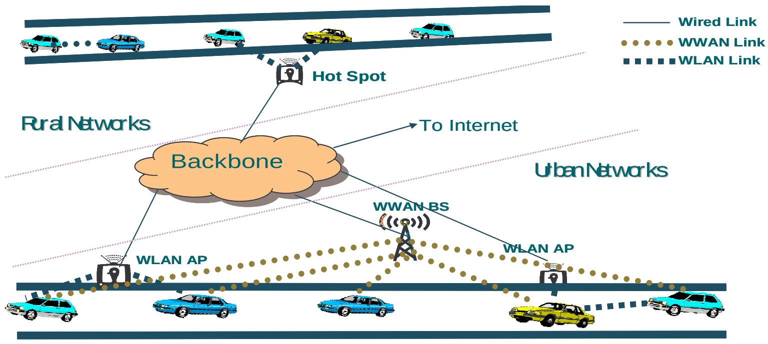


Figure 1: A notational vehicular network (note the V2V and V2R communications)

Figure 1 depicts a notational vehicular network one can easily envision in the near future. Multiple wireless technologies may coexist in an area. The rural and urban areas may deploy different network architectures. In urban areas wireless infrastructures such as cellular services provide nearly ubiquitous connectivity and Wi-Fi deployments continue to become more and more widespread. V2V communications can also be used for direct inter-vehicle information exchange. In rural areas, it might be more economical to rely on V2V communications supplemented by limited infrastructures placed in certain hot spots or other areas of particular interest. Figure 1 will be explained in more detail later.

Problem Statement: Vehicular communications are distinct from other types of wireless communication because of the high mobility of vehicles and the environment in which they operate. An understanding of the properties of different vehicular network architectures, e.g., delay, loss rate, throughput, traffic load, etc., is necessary before services can be successfully deployed. Based on this understanding, data services (e.g., dissemination services) tailored to specific vehicular network architecture can be designed to store, manipulate, aggregate, and transport data in an efficient way.

1.2 Infrastructure-less V2V Networks

A V2V network is an infrastructure-less network consisting only of instrumented vehicles. Vehicles are typically equipped with short-range communication devices. Vehicles can exchange information with other vehicles within their radio range, leading to the creation of ad-hoc wireless networks that can propagate information. This deployment offers the benefit of low cost and easy deployment, and is necessary for some localized applications (e.g., cooperative driving). A V2V network is a special type of ad-hoc network, exhibiting some unique characteristics [14, 78, 89]:

- Predictable high mobility - Vehicles often move at high speed, but on the road their mobility is rather regular and predictable. Vehicle movement is spatially constrained to the roadways, and vehicle operation is constrained by vehicle performance limitations and (at least to some extent) by traffic regulations, e.g., maximum and minimum speeds.
- Dynamic but geographically constrained topology - While the interconnection between vehicles can change quickly due to their high mobility the road network (which may be considered static) often limits the communication network topology to one dimension. A V2V network may be envisioned as overlaying the road network. Even where roads are in close proximity obstacles (e.g., buildings) generally prevent wireless signals from traveling between roads, except near intersections.
- Potentially large scale - Unlike most of the ad hoc networks studied in the literature that usually assume a limited area, V2V networks can in principle extend over the entire road network.

- Partitioned network - End-to-end connectivity is often implicitly assumed in ad-hoc networking research. However, Dousse et al. [21] showed that the probability of end-to-end connectivity decreases with distance for one-dimensional networks. Thus a V2V network is more likely to be partitioned, particularly at lower penetration ratios. This observation is also confirmed by analytical models [89] and simulation studies [90].
- Uncertain network reliability – Vehicles and in-vehicle devices are not completely reliable and dependable. They may fail in unpredictable ways.
- Instrumentation Capabilities – Vehicles can afford significant computing, communication and sensing capabilities. Mobile device power issues are usually not a significant constraint in vehicular networks as operating vehicles can provide continuous power to computing and communication devices.

The above characteristics have some implications on algorithm design on V2V networks. First, vehicle mobility can be exploited to propagate information from one network partition to another [89]. Secondly, due to the partitioned, highly dynamic nature of these networks, large-scale logical structures (e.g., trees) are undesirable; rather, localized algorithms [58] based on vehicles interacting with neighbors are preferred. Finally, unreliable communication channels, in-vehicle system failures, high mobility, and network partitioning introduce uncertainties in V2V networks. Data replication and diversity [30] can be employed to improve performance or increase data availability and reliability.

1.3 Infrastructure-based Vehicular Networks

In the preceding section, we have introduced infrastructure-less V2V networks. This deployment has many benefits but it fails to provide reliable communication services because it relies on unreliable V2V communications, especially where the density of instrumented vehicles is low. Also a pure V2V network as a standalone network cannot provide access to external online resources such as the Internet. It is often desired to offer infrastructure-based vehicular networks at least in some areas: to provide reliable broadband communication services, access online resources, communicate with other people, and access local services (e.g., traffic information, tourist information) not residing on vehicles. The infrastructure can provide two types of access: function-specific ports and communication ports. Vehicles communicate with the former for specific tasks. Examples include wireless-enabled intersection controllers enabling signal pre-empt (override for emergency vehicles) or signal priority (preferential treatment for mass transit vehicles), ramp meter controllers, and toll and parking payment collectors. Communication ports, e.g., WLAN access points (AP) and WWAN base stations (BS), provide network access, representing another type of access.

From the perspective of network infrastructure deployment, vehicular network infrastructures pose a number of distinctive characteristics. In addition to the mobility and instrumentation capabilities mentioned before, the distribution of vehicles deserves special attention. Conventional mobile users are often assumed to be concentrated in certain hot spots, e.g., buildings, airports, coffee shops, etc. However, vehicles are often widely distributed, where some roads (e.g., freeways) may have higher concentration of vehicles than others (neighborhood streets). On any specific road segment, the vehicle distribution may change dramatically due to accidents; in other cases, vehicle concentrations may be somewhat more predictable, as in the case of rush hours or congestion due to road work.

Traditionally, wireless infrastructures are invested heavily on hot spots (e.g., buildings, airports) where users are assumed to concentrate. However, this approach will not likely work for vehicular networks due to the wide distribution of vehicles. Deployment of wireless infrastructures for vehicles presents unprecedented challenges in financial cost, geographical scale and the number of users to be supported. From a commercial standpoint, such an infrastructure might be provided as a premium service to paid subscribers just as cellular and many WLAN services are provided today. Alternatively, services might be deployed by government, e.g., for traffic monitoring and management purposes, or economic development. Deployment of wireless infrastructures is one of the major focus areas of the ongoing Vehicle-Infrastructure Initiative, a joint undertaking between public (Federal, State, Local, Toll, etc.) and private(automotive companies, ITS equipment manufactures, communication companies, etc.) stakeholders [83, 84].

The infrastructure may leverage various wireless technologies, e.g., WWAN and WLAN, to work together in a seamless fashion. In the “urban network” of Figure 1, for example, WWAN base stations and WLAN access points are all connected to a backbone through wired links or fixed broadband wireless links (e.g., WiMAX [3]) which itself is connected to the Internet. Users can access the WWAN directly anywhere and anytime. However, due to its limited bandwidth and high cost, WWAN by itself may not be able to meet demands. For example, if a 2Mbps WWAN covers an area where there are 100 instrumented vehicles, a single vehicle can only obtain an average data rate of 2Kbps. This is clearly insufficient for delivering content-rich media. WLAN access points can be placed along the road to provide high-bandwidth and low-cost communication services. WLAN access points may exist independently or coexist with lights, poles, signs or emergency phones. They can also be set up on service areas, gas stations, or restaurants. A vehicle can access a roadside WLAN access point either directly or through the relay of other vehicles. A vehicle may be equipped with multiple wireless interfaces (e.g., cellular, 802.11x, DSRC etc) that attached to devices that are interconnected through internal network (e.g., Ethernet or Bluetooth). These devices cooperate to provide travelers the required communications services. This configuration effectively creates a floating "mobile intranet".

1.4 System Evaluation Methodology

System evaluation methodologies are essential to better understand vehicular networks and to design effective protocols and services. These evaluations must take vehicle traffic conditions, driving behavior, wireless communication characteristics, and protocol/application behavior into considerations. Statistical analysis, simulation, and field experiments provide alternate means for assessing system performance [28].

While on-road deployments provide the most realistic results, many vehicular network experiments are too difficult and/or dangerous to perform in a "live" setting or are cost-prohibitive as large-scale experiments. As a result, live experiments are usually conducted on a small scale. For example, Morsink et al. [59] demonstrated a co-operative collision warning and avoidance system to support longitudinal control of the vehicle using a fleet of 3 vehicles. Warnings based on acceleration, velocity and inter-vehicle headway data are generated to trigger both driver and automatic vehicle actions. In Chapter 5, we will present a set of experiments using proof-of-concept smart vehicles. Through these experiments, we investigate some basic operations in vehicular networks, e.g., V2V and V2R communications.

Understanding the behavior of large-scale systems requires the use of statistical analysis and/or simulation.

Statistical analysis [70] [89] offers a way to reach some basic understandings of system behavior and helps to derive meaningful settings for detailed evaluation. Statistical analysis often does not require very detailed data; rather coarse-grained statistical information is often adequate, making it especially attractive where detailed information is not readily available. Rudack et al. [70] derived mathematical models to study: (1) the duration when two vehicles stay within the range of each other; (2) the duration when one vehicle's neighbor list stays stable. However, statistical analysis normally employs simplifying assumptions (e.g., a Poisson process) to enable closed-form solution, making it less suitable to analyze complex systems. In Chapter 2 and 4, we will present some analytical models and conduct analysis using these models.

Simulation modeling can provide more realistic results and enable the study of detailed behaviors that are difficult or impossible to capture with analytical models. Combined simulation modeling includes simulation of vehicle movement, radio signal propagation, and protocol/application behavior. Often these models already exist separately, but need to be integrated to model vehicular networks. One approach is to develop a monolithic simulator incorporating all these models. This approach requires a significant amount work in development, modeling and verification. Another much less expensive approach is to leverage and extend existing diverse simulators, e.g., transportation [24] [82] [44] and wireless communication [72] [56] [16] simulators, enabling the re-use of models that have already been developed and verified. The problem then becomes one of integrating these simulators to interoperate and execute in a seamless fashion. The High Level Architecture (HLA) [71], a standard (IEEE 1516) developed by the U.S. Department of Defense has defined services for creating federated distributed simulation systems. Following this approach, we have developed a simulation-based test bed by federating two independent commercial simulation packages, QualNet [72] for communication simulation and CORSIM [24] for transportation system simulation. These two simulators were federated using a distributed simulation software package called the Federated Simulation Development Kit (FDK) [57] developed at Georgia Tech that provides services to exchange data and synchronize computations. FDK implements services defined in the Interface Specification of HLA. In addition to FDK, the test bed includes software developed for this research called the CORSIM-QualNet Communication Layer (CQCL) that not only defines interactions between CORSIM and QualNet, but also simplifies and streamlines the management of the distributed simulation execution. QualNet conducts a packet level telecommunication network simulation and implements the complete protocol stacks and physical environment for wireless communications. CORSIM is a microscopic traffic simulation model that simulates vehicle interaction, traffic flow, and congestion. The Run-Time Extension (RTE) facility available in CORSIM was utilized to extend the functionality necessary to operate the simulator in a distributed manner. For example, individual vehicle identification is retained when vehicles move between the freeway and arterial simulation modules. These unique vehicle IDs then flow from the traffic simulation to the communications simulation.

Instrumented vehicles in CORSIM are mapped to mobile nodes in QualNet to provide realistic mobility in the wireless network simulations. Common message formats are defined between CORSIM and QualNet for exchanging vehicle status and position information. During initialization, the transportation road network topology is transmitted to QualNet. Once the distributed simulation begins, vehicle position updates are sent to QualNet and are mapped to mobile nodes in the wireless simulation. Due to the large number of update messages, CQCL aggregates messages to reduce communication overhead. At the same time, the information arriving at mobile nodes in QualNet should also be sent to CORSIM as it may affect driving behavior, as in Ohio State University's OKI project [20], however, this has not been implemented at this time.

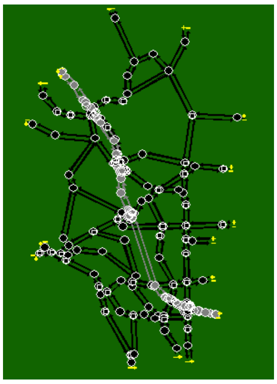


Figure 2: Vehicle traffic model simulating northwest quadrant of Atlanta, GA

Extensive geometric and operational data are required to model an area in CORSIM, including traffic flows, geometric layouts, intersection signal control parameters, observed vehicle speeds, travel times, etc. An extensive set of CORSIM models have been developed at Georgia Tech. For this research, we have frequently used a set of models simulating an area in the northwest quadrant of Atlanta, Georgia. These models incorporate approximately 12km of freeway and 160km of arterial surface streets, with approximately 5000 vehicles contained within the simulated region at any instant (Figure 2). Local and state government partners provided the majority of the required operational and geometric data, however, a calibration effort, including field surveys, was undertaken to insure that the CORSIM model provided a reasonable representation of actual operations. Moe detail can be found in the paper by Lee et al. [45].

1.5 Related Work

This section provides an overview of work and technologies in wireless networks, vehicular networks, and vehicle traffic theory to the extent that they are relevant to our research.

1.5.1 Wireless Networking

Wireless networking research has spanned a wide spectrum. Roughly speaking wireless networks can be classified as infrastructure-less (e.g., ad-hoc networks, sensor networks), infrastructure-based (e.g., cellular networks), or hybrid. For each type of network, there is an extensive body of research covering physical channel characterization, MAC protocol, transportation protocol, middleware, and application design. Here we review those most relevant to the research presented in this thesis.

1.5.1.1 Wireless Technologies

Depending on their coverage, wireless technologies can usually be categorized as Wireless Wide Area Networks (WWAN), Wireless Metro Area Networks (WMAN), Wireless Local Area Networks (WLAN), and Wireless Personal Area Networks (WPAN) in order of decreasing coverage. One inherent characteristic of all these technologies is that a larger coverage area often leads to higher cost per bit (see [68] for a cost comparison). For example, cellular-based WWAN often has large coverage areas (up to 20km), but offers relatively low bandwidth. The second generation (2G) systems (AMPS, TACS) provide less than 10Kbps of circuit switched data; 2.5G systems (GPRS, 1xRTT) offer less than 100Kbps packet switched data; 3G systems (WCDMA, UMTS) aim to offer data rates as high as 384Kbps outdoor and 2Mbps indoor. WMAN consists of many proprietary technologies and an ongoing standard – WiMAX (IEEE 802.16), currently providing fixed broadband wireless links. WLAN (e.g., IEEE 802.11x, HiperLan) is probably the most widely deployed wireless networks. They have limited cell coverage (~200m) but relatively high bandwidth (e.g., 802.11a/g provides data rates of up to 54Mbps). WPAN (e.g., Bluetooth, Zigbee) is primarily designed for networking personal devices and hence provides low transmission ranges (low power). WWAN typically operates in infrastructure mode with fixed-location base stations serving as access points, i.e., all wireless communications must go to or from base stations. WLAN can operate either in infrastructure mode or ad hoc mode. In ad hoc mode, mobiles relay packets for each other and no network infrastructure is needed, allowing the network to be readily deployed in environments such as battlefields and disaster relief sites where a fixed infrastructure does not exist, or has been damaged.

Today there is a major shift from backbones to edge networks. As a result we have seen a rapid growth and evolution of WWAN and WLAN technologies in providing network connectivity and Internet access to mobile users. While WWAN is aiming to provide ubiquitous coverage (e.g., Verizon Wireless Broadband Access), WLAN offers high bandwidth in a limited area (e.g., T-Mobile HotSpots). City-operated WLAN broadband networks are now under discussion in Chicago, Philadelphia, Las Vegas, New York, and San Francisco [2]. There is also extensive research in combining the two technologies to leverage the high capacity of WLAN and wide coverage of WWAN [33-35, 49, 52, 66, 92].

1.5.1.2 Information Theory

Information theory is a statistical theory dealing with the limits and efficiency of information processing. As stated in [5], research should be conducted to understand performance limits of wireless networks, with the goal of designing systems to reach these limits. Recent advances in information theory for wireless networks focus on the asymptotic sum throughput of a dense wireless network. Gupta and Kumar [31] proposed a model for studying the capacity of a static ad-hoc network. Their main result show that the throughput per session can be $O(1/\sqrt{n})$ at best, where n is the node density. Grossglauser and Tse [30] showed that a two-hop relay increases the per-session throughput capacity to $\Theta(1)$, i.e., independent of the node density. In traditional networking research, it is generally assumed that networks are connected. The end-to-end delay is determined by the number of hops and the queuing and processing delay in each hop. However, this is not the case for a partitioned network (e.g., V2V networks).

The connectivity probability for any two nodes in one-dimensional and two-dimensional areas is presented in [21]. Cheng and Robertazzi [18] analyzed the coverage of broadcast in multi-hop radio networks. Some of our statistical analysis to be presented later is based on these works.

1.5.1.3 Ad-Hoc Networks

Conventional practices in the area of wireless ad-hoc networks focus on the network in a small two-dimensional area. Most studies assume mobile nodes move randomly and any two nodes can be expected to be close to each other from time to time due to their confinement to a limited area [32, 42, 64], or alternatively, that topology information can be maintained with low cost due to slow or no node mobility [37, 76]. In a vehicular environment, due to high mobility, maintaining inter-vehicle connection based topology will typically be expensive, possibly even infeasible, and likely unnecessary. It is also not reasonable to assume that two encountering vehicles will meet again with any certainty except some particular fixed-schedule vehicle systems (e.g., transit systems) or commuting vehicles.

The concept of a partitioned network exploiting node movement to deliver information opportunistically is well explored. In [19, 81], mobile hosts exchange information when they meet. In [48], mobile nodes proactively change their movement to deliver messages. Message Ferry [97] [98] provides a common communication channel via a node of a known fixed moving trajectory for disconnected mobile nodes. The Zebra project [40] and the MULE architecture [73] are intended to provide intermittent connectivity in a disconnected ad-hoc network. Fall [23] proposed a framework for supporting delay tolerant networks (DTN). One such application is the DakNet project [65], which provides information services to remote rural villages. Compared to these works, we need to specifically address vehicle mobility.

Flooding is expected to be performed frequently in a vehicular environment due to both the high system dynamics and peculiar applications to be deployed (e.g., safety applications, information services). The broadcast storm problem arises when disseminating information throughout the network where serious redundancy, contention, and collision could exist. The fundamental approach to solve the broadcast storm problem is to inhibit some nodes from re-transmission [61] [76]. These techniques can be adapted in vehicular networks.

Most mobility models, as in [96], allow nodes to move randomly. These models are not applicable to vehicles because vehicles must stay on the road. In the Manhattan mobility model [13], nodes are moving along streets, but wireless signals can still travel across streets as if there is no building between them. This is obviously not realistic. In our research, instead of synthesizing some mobility model, we let the node movement be driven by the microscopic transportation simulation.

1.5.1.4 Data Delivery Mechanism

Data delivery mechanisms define the rules for moving information throughout the network. Conventional data delivery services often implicitly assume that the network is connected. The “node centric” approach [38] specifies the routing path as a sequence of connected nodes. However, high node mobility will quickly render inter-node connections invalid. “Geographical forwarding” [54] can accommodate high mobility by decoupling the routing path from the intermediate nodes. The message is forwarded by selecting the next hop(s) such that the message will be moved closer to the geographic destination. If a gap in the sought direction is encountered, efforts are made to find a path around it [41]. This approach will fail when the network is partitioned (or at least non-continuously connected) and no immediate end-to-end path is available.

“Opportunistic forwarding”, as suggested in [17, 23], targets networks where an end-to-end path cannot be assumed to exist. Messages are stored and carried by mobile nodes and forwarded as opportunities present themselves. When a message is forwarded to another node, a copy may remain with the original and can be forwarded again later to improve reliability or performance. Simple opportunistic forwarding does not consider the final message destination, passing a message whenever possible, e.g., two nodes exchange data whenever they can communicate [32, 42, 64]. It is effective for transmitting data to all potential receivers, but inefficient if a message is to be delivered to only specific receivers, e.g., those in a certain region. In this case, overall network performance may be improved if messages are only forwarded such that they migrate closer to their eventual destination, and not unnecessarily consume network resources in other areas.

"Trajectory based forwarding" [62] directs messages along predefined trajectories. It was originally presented to work well in a dense network. However, it should also work on graph-based networks, e.g., V2V networks, because both nodes and messages are moving along graph edges. Trajectory based forwarding can help limit data propagation and thus reduce message overhead.

1.5.2 Vehicular Networks

Studies concerning vehicular networks have largely focused on ad-hoc V2V communications. The Fleetnet project [25] and its successor Network-On-Wheels (NOW) project, and California PATH project [15] has investigated radio devices [51], MAC protocols [94] and routing protocols [79] [50] for V2V communications.

Most safety-related services proposed so far use one hop communications. In particular, single-hop broadcast is often proposed for safety applications where the focus is on improving reception reliability [80, 93, 95]. For example, a braking vehicle can notify the vehicles behind its intention to avoid possibly hazardous conditions. On the other hand, multi-hop communications can be utilized to extend communication range [87, 90]. For instance, the coverage of a roadside service station can be substantially extended through

V2V communications. Multi-hop V2V communications may be a viable approach to provide communication services over a long distance, particularly where cellular access is not available or too expensive. Due to the high mobility inherent in V2V networks, most routing protocols proposed use position-based addressing and forwarding [54], and assume end-to-end connectivity [79]. An adaptive information propagation scheme is introduced in [85].

There are a few papers [43] [63] talking about placing WLAN-based Internet gateways along the road. For example, Kutzner et al. [43] discussed how to leverage wired gateways along the road to assist routing in a hybrid network structure. However no detailed implementation was given.

There has been some initial work in measuring the performance of V2V and V2R communications. Singh et al. [74] measured the performance of V2V 802.11b communications. They categorized operating environments as suburban, urban and freeway, and reported the communication performance in each category. Ott and Kutscher [63] investigated the TCP/UDP performance of a car driving by a roadside WLAN access point. The above efforts focus on single-hop communications. Möske et al. [60] demonstrated a mobile multi-hop scenario. These studies have explored the impact of various communication parameters (e.g., packet size, data rate, signal level), however, they have not attempted to relate the vehicle operating environment to communication performance.

1.5.3 Vehicle Traffic Flow Theory

Vehicular networks consist of communication devices residing inside vehicles. Vehicle traffic conditions naturally play an important role. Vehicle traffic is traditionally modeled using vehicle traffic flow theory [29] [55] in the transportation research literature. Next, we introduce some of the basics of relevant vehicle traffic flow theory.

The three fundamental characteristics of vehicle traffic are flow q (vehicles/hour), speed u (km/h) and density λ (vehicles/km). The average values of these quantities can be approximately related by the basic traffic stream model $u = q / \lambda$. With few vehicles on the roadway, density approaches zero and speeds approach free flow speed. As additional vehicles enter the roadway traffic density and flow increase, until flow reaches a maximum. As additional vehicles continue to enter the traffic stream density will continue to increase but the flow will begin to decrease. As demand exceeds roadway capacity, increasing traffic densities approach a "jam density" limit. At this limit all vehicles are stopped (i.e. flow is zero) with vehicles tightly packed on the roadway, i.e. a traffic jam. This model also holds if a subset of vehicles is considered (e.g., instrumented vehicles).

The actual relationships are much more complex, but can be readily modeled in simulation models [24] [82] [44]. These and other simulation models employ car following theory to implement vehicle-vehicle interactions and track the motion of individual or platoons of vehicles. There are in general two types of vehicle traffic models: macroscopic and microscopic models. Macroscopic models study the vehicle traffic flow as a whole (typically, fluid flow models are used) while microscopic models study individual vehicle behavior. Microscopic models are of greater interest in studying vehicular networks because they allow the capture of inter-vehicle interaction and the continuously changing distribution of instrumented vehicles among the traffic stream. In cases where simulation models cannot be employed due to a lack of data, flows can be characterized with discrete distributions such as Poisson (for low-density traffic), negative binomial (for varying flow), and binomial distribution (for congested flow); time headways between the arrivals of vehicles may be represented with an exponential, shifted exponential, Erlang, or normal distribution; the distribution models for speeds are usually normal or lognormal.

Traffic monitoring is commonplace in most major urban areas. Given the importance of minimizing congestion, major metropolitan areas expend millions of dollars per year to monitor freeway speeds and flows in an effort to identify and remove disabled vehicles from the roadway. One example is Traffic Management Center (TMC) of Georgia Department of Transportation (GDOT).

1.6 Research Contributions

In this research, we have explored several issues concerning the analysis and design of vehicular networks. This study is not intended to cover every aspect of vehicular networks, but rather serve to provide insights into certain aspects of vehicular networks as well as to motivate additional work in this area.

- *Analysis of vehicle-to-vehicle (V2V) communications*: V2V communications are an important component in many vehicular networks. However, V2V communications have certain limitations. In particular, we explore the spatial propagation of information using V2V communications. Analytical models based on simplified assumptions are developed to give a basic understanding. Simulations using realistic vehicle traffic models are used to study more complex, realistic vehicle traffic scenarios.
- *Design of enhanced opportunistic forwarding algorithms*: Based on the above analysis, opportunistic forwarding appears to be a viable approach for data

dissemination using V2V communications for applications that can tolerate some data loss and delay by exploiting vehicle mobility to overcome vehicular network partitioning. We present a generic methodology to design enhanced opportunistic forwarding algorithms. Two algorithms, MDDV and *optimistic forwarding*, are derived out of the generic methodology.

- *Architecture study of infrastructure-based vehicular networks*: Several architecture alternatives for infrastructure-based vehicular networks have been identified and evaluated in a realistic vehicular environment. The evaluation reveals certain characteristics of the main vehicular network infrastructure building blocks and their interactions. The evaluation results lead to some insights concerning issues such as the use of multi-hop forwarding in vehicular networks, as well as the approaches to address varying vehicle traffic conditions.
- *A detailed measurement study of short-range communications between vehicles and between vehicles and roadside stations on a highway*: We conduct field experiments to study the expected wireless communication characteristics in a driving environment and identify factors that significantly affect communication performance. We also demonstrate the benefits of multi-hop communication in improving communication performance.

1.7 Roadmap for This Thesis

The remainder of the thesis is organized as follows. Chapter 2 discusses the spatial propagation of information using V2V communications. In Chapter 3, we present the design of enhanced opportunistic forwarding algorithms for data dissemination using V2V communications. In Chapter 4, we identify and evaluate several architecture alternatives for infrastructure-based vehicular networks. Chapter 5 presents a measurement study of short-range communications. Chapter 6 summarizes this thesis and suggests the future research directions.

CHAPTER

2 SPATIAL PROPAGATION OF INFORMATION

2.1 Introduction

An understanding of properties of vehicular networks, e.g. delay, loss rate, throughput, etc., is necessary before proposed applications can be successfully deployed. As stated in [5], information theory research should be conducted to understand performance limits of wireless networks, with the goal of designing systems to reach these limits. An area of increasing interest concerns the use of *short-range vehicle-to-vehicle (V2V)* communications rather than relying completely on roadside infrastructures that are expensive to deploy and maintain. This is the type of system examined in this chapter. A V2V network is a special type of ad-hoc networks. An overview of V2V networks was provided in Chapter 1.2. The end-to-end (E2E) delay due to network partitioning is one of the major limitations of V2V networks. In this chapter, we study the spatial propagation of information: *how fast can information propagate along a specific road?* Stated another way, what is the average delay to propagate a message from location A to location B separated by a distance x along a specific road? The answer to this question can help address issues such as how far a piece of traffic information can propagate before becoming obsolete. It also enables an in-depth understanding of the limitations of V2V networks and whether additional communication facilities such as roadside base stations will be required to support specific applications.

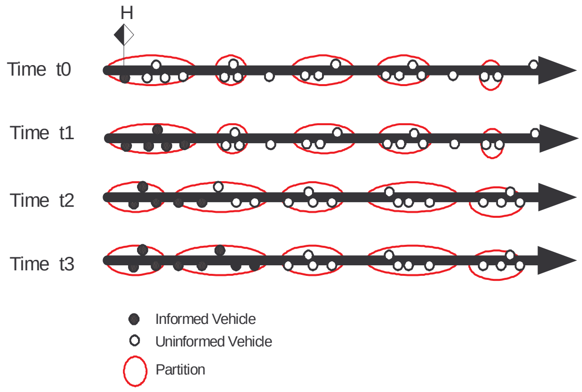
In traditional networking research, it is generally assumed that networks are connected. The end-to-end delay is determined by the number of hops and the queuing and processing delay in each hop. However, V2V networks are partitioned. Prior research [17] [90] has shown that the motion of vehicles can contribute to message delivery. We thus focus on how vehicle mobility patterns impact information propagation.

We use both statistical analysis and simulation to approach this problem. We first present analytical models based on simplified assumptions to provide a basis. Statistical analysis only requires some coarse-grained statistical information (e.g., vehicle traffic flow rate and speed distribution) as input. We then introduce simulation studies using realistic vehicle traffic models incorporating detailed transportation system operation data.

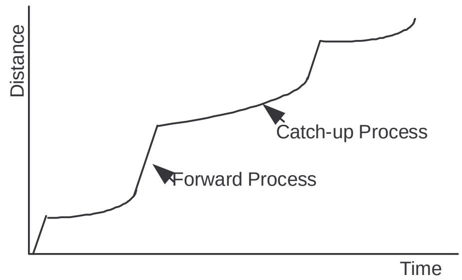
2.2 Analytical Models

2.2.1 Definitions and Assumptions

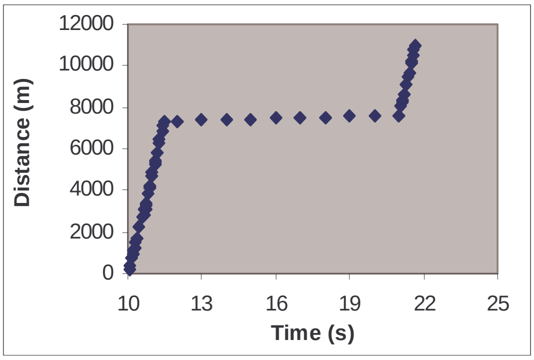
To analyze information propagation speed, we employ an idealized data propagation model. We call a vehicle *informed* if it carries the message being propagated. When an uninformed vehicle enters the radio range of an informed one, it becomes immediately informed. Every instrumented vehicle is assumed to have the same radio range r. A vehicle requires a fixed amount of time tr to receive and process a message before the message is available for further retransmission. In this way, our analysis is neutral to specific wireless technologies, e.g., 802.11x, HiperLan etc. Here we do not consider many real-world communication aspects, e.g., signal interference, bandwidth constraint, and link quality variation, etc. This idealized communication scheme allows us to simplify the analysis and determine an upper bound on propagation speed.



(a)



(b). Spatial-Temporal Trajectory



(c). Information Propagation Trace

Figure 3: An example of message propagation

We study information propagation in one direction along the road. Figure 3(a) shows an example. The arrow represents the road and the circles below (above) the arrow represent vehicles traveling in the positive (negative) direction. The vehicle traffic pattern along the road is assumed to be statistically uniform along the entire roadway. Since the typical radio range is much larger than the road width, only one-dimensional distance along the road is considered; the road’s width is neglected. Without loss of generality, we consider a message propagating in the positive direction from a vehicle at location H at time 0. Position coordinates of vehicles increase to the right. The *message head* at time t refers to the informed vehicle with the largest position coordinate, representing the farthest point to which the message has propagated. The *partition tail* at time t refers to the uninformed vehicle with the smallest position coordinate to the right of location H. Note that the message head and partition tail vehicle may change over time.

The process of message propagation is the same as the traveling process of the message head. At time t_1 in Figure 3(a), the message reaches the front most vehicle of its current partition through multi-hop forwarding, and begins to travel with the message head. The message cannot reach the next partition at this time because the gap between the foremost instrumented vehicle in the first partition and the rearmost instrumented vehicle in the second partition is greater than the radio range. At time t_2, the message head catches the next partition through the relative movement of vehicles. At time t_3, the message traverses the next partition and then the above process repeats again. A notational spatial-temporal trajectory of message propagation is presented in Figure 3(b). Figure 3(c) illustrates a simulated message propagation trace along I-75 of Northwest Atlanta, GA. A message begins to propagate at 10s. At 10.48s, the message reaches the head of its current partition and starts to move with the message head at the speed of this particular vehicle until 21s when the message reaches the radio range of another partition. The simulation experiments and results will be described in greater detail later.

From the above examples we make the following observations:

- The V2V network is often a partitioned network. A snapshot of the position of vehicles at any time instant will typically yield a network with many partitions, as is the case in Figure 3(a). Vehicles within a partition can communicate (perhaps requiring several hops), but no direct connection exists between partitions.

- The configuration of partitions is dynamic. Partitions may split or merge as relative positions of vehicles within the partition or between partitions change.
- A message propagates in either one of two processes termed the *forward process* and *catch-up process* as shown in Figure 3(b). The forward process involves the propagation of the message within a partition via multi-hop forwarding; the message travels quickly through a partition hop-by-hop until it reaches the front most vehicle in that partition. In the catch-up process the message moves along with its carrying vehicle until it comes within the radio range of the last uninformed vehicle in the partition ahead of it. Due to limited vehicle speed, the information propagation speed in the catch-up process will normally be much slower than that in the forward process. When information propagates along a partitioned road network, it alternates between the forward process and catch-up process, resulting in a cyclic process.

Before we study information propagation, we must define the vehicle traffic model. An undisturbed vehicle traffic model is used here with the following assumptions:

- Poisson arrival. Traffic passing an arbitrary point on the road follows a Poisson process with an average rate equal to the traffic flow rate (vehicles/unit time).
- Random and independent vehicle mobility. Each vehicle travels with an average speed that is selected from a random distribution in the interval [$v_{min,}$ v_{max}]. Vehicles move independently at their chosen velocity.

This model is utilized to enable closed-form solution of the quantities of interest. This model is only applicable when traffic is not congested and there are no disturbing factors such as traffic signals, so is more applicable to uncongested freeway traffic. This model, though simplified, captures the basic dynamics of vehicle interactions and allows one to examine important vehicle traffic characteristics that significantly influence the rate of information propagation. Later, the validity of the model as applied to message propagation is examined through comparisons with results obtained through using detailed vehicle activity on a highway. Vehicle mobility has been widely studied in transportation research [29, 55] . However close-form solutions for information propagation are difficult to obtain for more complex models, making them better suited for simulation analyses (Section 2.4).

In this section we explore message propagation with one-way vehicle traffic. We are primarily interested in the scenario where the message propagates in the direction in which vehicle traffic is traveling though later we will show that messages may also propagate along vehicles that are traveling in the opposite direction of message flow because radio transmission speed is much faster than vehicle speed.

2.2.2 Qualitative Analysis

The forward process is straightforward, while the catch-up process is more complex. Here we give out a detailed analysis of the catch-up process. A message may not be able to catch the partition in front of it. In the one-way vehicle traffic scenario, it is easy to see that if all the vehicles are moving at the same speed, the relative positions of vehicles and thus inter-vehicle gaps remain static, and a message can never advance beyond a partition. Therefore relative vehicle movement (either in the same direction or opposite direction) is necessary for inter-partition message propagation to occur.

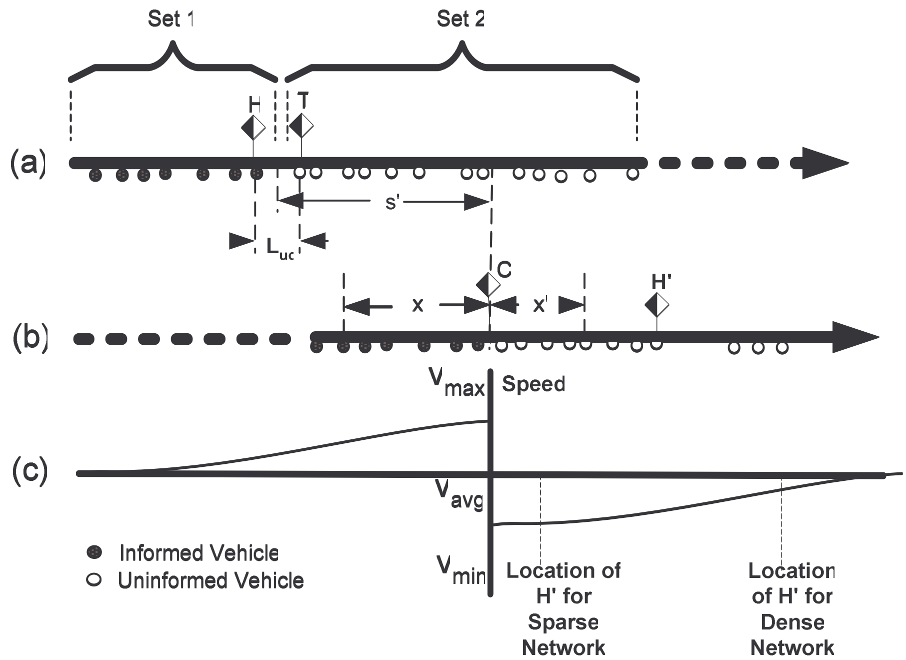


Figure 4: Catch-up process

Figure 4 illustrates a catch-up process. Figure 4(a) shows the beginning of the catch-up process. H is the location of the message head. Vehicles are in either one of two sets: informed vehicles are in set 1, and uninformed vehicles are in set 2. Figure 4(b) shows the end of the catch-up process when the message head just enters the radio range of the last uninformed vehicle in set 2. Location C is the catch-up point (the midpoint between the message head and the rearmost uninformed vehicle). H' is the location of the front most vehicle of the partition caught and will be the location of the message head in the beginning of the next catch-up process if we ignore the time spent in intra-partition forwarding (which is short relative to the time required for relative vehicle movements). During the catch-up process, set 1 and set 2 do not interleave. Figure 4(c) shows the temporary non-uniform distribution of average vehicle speed around C in the end of the catch-up process. During the catch-up process, fast informed vehicles in set 1 are likely to overtake the slow informed vehicles and concentrate near C, and the slow vehicles in set 2 tend to lag behind and stay close to C. This illustrates that non-uniform vehicle traffic distribution can occur in a microscopic scale even though the vehicle traffic is assumed to be uniform at the macroscopic level. We develop an approximate analysis to further demonstrate this.

Suppose the catch-up process takes a time period of t. We want to compute $E[V_t(x)]$, the average speed of informed vehicles with distance x from C in set 1, and $E[V_t(x')]$, the average speed of uninformed vehicles with distance x' from C in set 2, in the end of the catch-up process (as shown in Figure 4(b)). We have

$$V_t(x') \in [v_{min}, \min(\frac{x'+s'}{t}, v_{max})] \quad x' > 0$$

Suppose $V_t(x')$ is uniformly distributed in this speed interval. If $\frac{x'+s'}{t} < v_{max}$, $E[V_t(x')]$ will be less than the average speed among all vehicles. While x' decreases, i.e., the studied location is closer to C, $E[V_t(x')]$ also decreases. Similar analysis shows that $E[V_t(x)]$ increases when x decreases.

In general, two consecutive forward and catch-up processes are not independent because some vehicles may participate in both processes, making it complicated to analyze the entire process. However two extreme cases are relatively simple to explore. If the network is sparse in the sense of low instrumented vehicle density (with respect to the vehicle radio range), the partition will be short and H' is very close to C in Figure 4(b). During the next catch-up process, it is very likely that some fast vehicle from set 1 is able to overtake the slow message head (from set 2) as a result of the non-uniform speed distribution around C in the end of the current catch-up process (Figure 4(c)). This situation is more likely to happen with sparser networks because partitions are smaller.

This observation motivates our models for sparse networks to be addressed in Chapter 2.2.4. On the other hand, if the network is dense, H' is far from C. We can ignore the non-uniform speed distribution around C and assume that vehicle traffic probabilistically restarts in the beginning of each catch-up process. This observation leads to our models for dense networks to be introduced in Chapter 2.2.5.

2.2.3 Notation

We summarize the notations used to define our models in Table 1. Capitalized symbols denote random variables.

Table 1: Notations and Model Parameters

P[e]	Probability of an event e
E[R]	Expectation of a random variable R
$F_R(r)$	CDF of a random variable R
$f_R(r)$	PDF of a random variable R
N(t)	Number of vehicles passing location H, the location of the message head at time 0, during (0, t]
X(t)	Message propagation distance during [0, t]
V(t)	Message propagation speed at time t
N′(t)	Number of vehicles passing location T, the location of the last uninformed vehicle at time 0, during $[-(v_{max} / v_{min} - 1)\ t, 0)$
X′(t)	Distance that the partition tail moves during [0,t]
V′(t)	Speed that the partition tail moves at time t
V_i	Average speed of vehicle i, i = 0, 1…n
T_i	Time when vehicle i passes either location H or location T, i=1, 2…n
T_c	Time of a catch-up phase
X_c	Message propagation distance during a catch-up phase
T_f	Time of a forward phase
X_f	Message propagation distance during a forward phase
Y	Partition size
M	Number of vehicle gaps in a partition
L	Gap between two neighboring vehicles
L_c	Gap between two neighboring connected vehicles
L_{uc}	Gap between two neighboring disconnected vehicles
λ	Traffic flow rate (vehicles/unit time)
μ	Density of instrumented vehicles (vehicles/unit distance)
r	Vehicle radio range
tr	Message processing time

2.2.4 *Sparse Network Model*

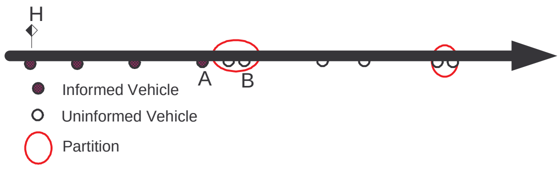


Figure 5: "Re-overtake" scenario

In a sparse network, the density of instrumented vehicles is sufficiently low so that the vehicle radio range is quite small compared to the average inter-vehicle gap. As a result the average partition size is small. We illustrate the information propagation process with the example shown in Figure 5. Vehicle A is the message head that just catches the partition ahead of it. The message quickly propagates through the new partition, and vehicle B now becomes the new message head. However, as described in Chapter 2.2.2, it is quite likely that vehicle A (or some faster informed vehicle) will eventually pass slow vehicle B later and become the message head again. We call this the "re-overtake" scenario. We can thus approximate the message propagation as depending entirely on vehicle movement rather than on intra-partition forwarding, i.e., ignoring the message forwarding within partitions. This is equivalent to the scenario that the vehicle radio range is close to 0 so that two vehicles only communicate when passing each other. We use this equivalent scenario when deriving our sparse network model. These assumptions hold when most partitions only have one vehicle except when one vehicle overtakes another vehicle (when the average distance between two neighboring vehicles is much larger than the radio range).

Consider a message originating from a vehicle of speed V_0 residing at location H at time 0 in Figure 5. Let us define N(t) to be the number of vehicles passing location H during (0, t]. Based on our assumption that the radio range is close to 0, only vehicles passing location H after time 0 can be the message head and thus the message head at time t is the vehicle of the maximum position coordinate among those vehicles passing location H during [0, t]. X(t), the traveling distance of the message head during [0,t], is

$$X(t) = \max(V_0 * t, \max(V_i * (t - T_i))) \quad i = 1, 2, ..., N(t) \qquad (2\text{-}1)$$

where V_i is the average speed of vehicle i, and T_i is the time when vehicle i passes location H after time 0. Next we derive the distribution of X(t). Conditioned on N(t), we have

$$F_{X(t)}(x) = \sum_{n=0}^{\infty} P[X(t) < x \mid N(t) = n]\, P[N(t) = n] \qquad (2\text{-}2)$$

Following (2-1), we have

$$\begin{aligned} &P[X(t) < x \mid N(t) = n] = \\ &P[V_0 * t < x, V_i * (t - T_i) < x \ \text{ for each } \ i = 1, 2, ..., n] \end{aligned} \qquad (2\text{-}3)$$

By the assumption of undisturbed vehicle traffic model, vehicles passing location H follow a Poisson process. It can be shown (e.g., see Theorem 5.2 of [69]) that, under the condition that n vehicles pass location H during (0, t], the times at which these vehicles pass location H, considered as unordered random variables, are distributed independently and uniformly in the interval (0, t]. This leads to

$$T_i \sim \text{uniform}(0,\ t) \qquad (2\text{-}4)$$

Since T_1, T_2…T_n and V_0, V_1…V_n are i.i.d, we can drop the subscripts

$$P[X(t) < x \mid N(t) = n] = P[V * t < x] * P[V * (t - T) < x\,]^n \qquad (2\text{-}5)$$

where T has the same distribution as T_i and V has the same distribution as V_i.

Based on the assumption of Poisson process, N(t) has the Poisson distribution

$$P[N(t)=n]=\frac{e^{-\lambda * t} *(\lambda * t)^n}{n!} \qquad (2\text{-}6)$$

where λ is the traffic flow rate (vehicles/unit time). Substitution of (2-5) and (2-6) in (2-2) yields

$$F_{X(t)}(x)=P[V*t<x]*\sum_{n=0}^{\infty}\left(P[V*(t-T)<x]^n*\frac{e^{-\lambda * t}(\lambda * t)^n}{n!}\right) \qquad (2\text{-}7)$$

Further we have

$$f_{X(t)}(x)=\frac{\partial F_{X(t)}(x)}{\partial x} \qquad (2\text{-}8)$$

$$E[X(t)]=\int_{-\infty}^{\infty} x*f_{X(t)}(x)dx \qquad (2\text{-}9)$$

Let V(t) be the message head speed,

$$V(t)=\frac{\partial X(t)}{\partial t} \qquad (2\text{-}10)$$

$$E[V(t)]=\frac{\partial [E[X(t)]}{\partial t} \qquad (2\text{-}11)$$

Thus given the distribution of vehicle speed V and the vehicle flow rate λ, we can obtain the distribution of X(t) in equation (2-7). E[V(t)], the average message head speed, is given in equation (2-11). We show a numerical computation of E[V(t)] in Figure 6. All numerical results were generated using *mathematica* [86]. The results shown in Figure 6 are obtained by assuming the vehicle speed distribution is V ~ uniform(v_{min}, v_{max}), where v_{min} = 11.11m/s and v_{max} = 20m/s, and the traffic flow rate λ = 3600vph (these settings can be used to generate a sparse network when the vehicle radio range is close to zero). It is clear from the result that on average, message propagation accelerates from the average speed of all vehicles and is bounded by the maximum vehicle speed. Intuitively, this reflects the fact that if there is a speed discrepancy between vehicles, when the message head is overtaken by another vehicle from behind, the overtaker becomes the new message head and has a higher speed than the previous message head. The message propagation speed thus tends to increase, and the message head migrates to faster vehicles over time.

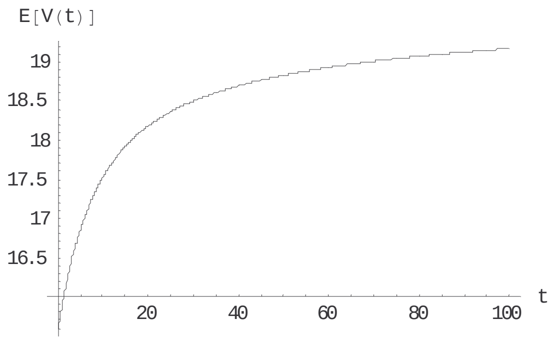


Figure 6: Average message head speed

2.2.5 Dense Network Model

In a dense network, the average distance between two neighboring vehicles is similar to or smaller than, the vehicle radio range. As described in Chapter 2.2.2, it is a reasonable approximation to assume two consecutive catch-up processes are independent since they are separated by a large partition, and thus vehicle traffic probabilistically restarts in the beginning of each catch-up phase. For a given time t, we define $V_p(t)$ to be the average message propagation speed during the interval [0, t], $V_p(t) = X(t) / t$. We define the long-term average message propagation speed v_p to be

$$v_p = \lim_{t \to \infty} V_p(t) = \lim_{t \to \infty} \frac{X(t)}{t} \qquad (2\text{-}12)$$

The message propagation process can be modeled as a *renewal reward process* with a new cycle beginning each time the catch-up process begins and the reward is the message propagation distance during each cycle. Each cycle consists of a catch-up phase followed by a forward phase. It can be shown (e.g., see Chapter 7.4 of [69]) that

$$v_p = \frac{E[X_c + X_f]}{E[T_c + T_f]} = \frac{E[X_c] + E[X_f]}{E[T_c] + E[T_f]} \qquad (2\text{-}13)$$

where X_c (X_f) is the average distance traveled during the catch-up (forward) phase and T_c (T_f) is the time spent in the catch-up (forward) phase. Next, we will examine these two phases respectively.

2.2.5.1 Catch-up Phase

We want to compute $E[T_c]$ and $E[X_c]$. During the catch-up phase, the message head will eventually catch up to the partition tail. We need to analyze the movement of both the message head and partition tail.

The beginning of a catch-up phase is illustrated in Figure 4(a). Let us assume a cycle begins at time 0 and the locations of the message head and partition tail at time 0 are H and T, respectively. Through the same analysis as in 2.2.4, we can obtain the distribution of X(t), the distance that the message head moves during [0,t]. The result is given in equation (2-7).

Now we examine the movement of the partition tail. Note that the partition tail is the last uninformed vehicle in the partition ahead and the vehicle taking this role may change over time. By assumption the vehicles passing location T form a Poisson process. No uninformed vehicle passes location T after time 0 during the catch-up phase. Define V_0 to be the speed of the vehicle at location T at time 0, and N'(t) the number of vehicles passing location T during [-(v_{max}/ v_{min}-1) * t, 0). The partition tail at time t is the vehicle with the minimum position coordinate among all vehicles passing location T during the time interval [-(v_{max}/ v_{min}-1) * t, 0]. Let X'(t) be the distance that the partition tail moves during [0,t], we have

$$X'(t) = \min(V_0 * t, \min(V_i * (t - T_i))) \quad i = 1, 2, ..., N'(t) \tag{2-14}$$

while T_i is the time when vehicle i passes location T ($T_i < 0$).

Conditioned on N'(t), we have

$$F_{X'(t)}(x) = 1 - \sum_{n=0}^{\infty} P[X'(t) > x \mid N'(t) = n] * P[N'(t) = n] \tag{2-15}$$

Following equation (2-14), we have

$$\begin{aligned} &P[X'(t) > x \mid N'(t) = n] = \\ &P[V_0 * t > x, V_i * (t - T_i) > x \text{ for each } i = 1, 2, ..., n] \end{aligned} \tag{2-16}$$

Based on the properties of Poisson process

$$T_i \sim \text{uniform}(-(v_{max} / v_{min} - 1) * t,\ 0) \tag{2-17}$$

$$P[N'(t) = n] = \frac{e^{-\lambda*(v_{max}/v_{min} - 1)*t} * (\lambda * (v_{max} / v_{min} - 1) * t)^n}{n!} \tag{2-18}$$

Again based on the properties of i.i.d

$$P[X'(t) > x \mid N'(t) = n] = P[V * t > x] * P[V * (t - T) > x]^n \tag{2-19}$$

where T has the same distribution as T_i and V has the same distribution as V_i.

We can derive $F_{X'(t)}(x)$ by substituting equation (2-18) and (2-19) into (2-15).

Further we have

$$f_{X'(t\}}(x) = \frac{\partial F_{X'(t)}(x)}{\partial x} \tag{2-20}$$

Let V'(t) be the partition tail speed,

$$E(X'(t)) = \int_{-\infty}^{\infty} x * f_{X'(t)}(x)\, dx \tag{2-21}$$

$$E[V'(t)] = \frac{\partial [E[X'(t)]]}{\partial t} \tag{2-22}$$

Thus given the distribution of vehicle speed V and the vehicle flow rate λ, we can derive the distribution of X'(t) in equation (2-15). E[V'(t)], the average partition tail speed, is given in equation (2-22). A numerical result of E[V'(t)] is shown in Figure 7. The system parameters are set the same as in Chapter 2.2.4. It is clear that on average the partition tail decelerates from the average speed among all vehicles and is bounded by the minimum vehicle speed. Intuitively, this shows that when the partition tail passes another vehicle from behind, the passed vehicle becomes the new partition tail and is traveling more slowly than the previous one, i.e., the partition tail migrates to slower vehicles over time.

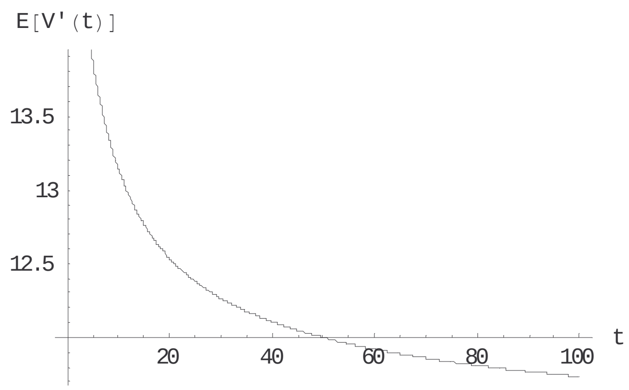


Figure 7: Average partition tail speed

So far we have shown that the message head is accelerating while the partition tail is decelerating so that eventually the message head can catch up to the partition that lies ahead. This verifies the analysis in Chapter 2.2.2.

Next, we derive the distribution of the catch-up time T_c. Conditioned on L_{uc}, the gap between location H and T in Figure 4(a), we have

$$\begin{aligned} P[T_c \le t \mid L_{uc} = l] &= P[X'(t) \le X(t) + r - l] \\ &= \int_0^{\infty} f_{X(t)}(x) \int_0^{x+r-l} f_{X'(t)}(x')\, dx'\, dx \qquad l > r \end{aligned} \tag{2-23}$$

where r is the vehicle radio range. This yields the distribution of T_c (t)

$$F_{T_c}(t) = P[T_c \le t] = \int_r^{\infty} P[T_c \le t \mid L_{uc} = l] * f_{L_{uc}}(l)\, dl \qquad (2\text{-}24)$$

Two neighboring vehicles are disconnected if their distance is larger than r. Given the distribution of the vehicle gap L, we have

$$f_{L_{uc}}(l) = f_L(l \mid L > r) = \begin{cases} \dfrac{f_L(l)}{1 - F_L(r)} & l > r \\ 0 & \text{else} \end{cases} \qquad (2\text{-}25)$$

We can derive F_{T_c}(t) by substituting equation (2-23) and (2-25) into (2-24).

Again

$$f_{T_c}(t) = \frac{\partial F_{T_c}(t)}{\partial t} \qquad (2\text{-}26)$$

$$E[T_c] = \int_0^{\infty} t * f_{T_c}(t)\, dt \qquad (2\text{-}27)$$

$E[X_c]$ can be computed using f_{Tc} (t)

$$\begin{aligned} E[X_c] &= \int_0^{\infty} E[X_c \mid T_c = t] * f_{T_c}(t)\, dt \\ &= \int_0^{\infty} E[X(t)] * f_{T_c}(t)\, dt \end{aligned} \qquad (2\text{-}28)$$

where E[X(t)] can be obtained from equation (2-9).

Finally equation (2-27) presents $E[T_c]$ and equation (2-28) gives out $E[X_c]$.

2.2.5.2 Forward Phase

We now compute $E[T_f]$ and $E[X_f]$. Let Y be the size of a partition. We have

$$E[X_f] = E[Y] + r \qquad (2\text{-}29)$$

$$E[T_f] = \left\lceil \frac{E[X_f]}{r} \right\rceil * tr \qquad (2\text{-}30)$$

Thus we only need to determine the average partition size Y. Two neighboring vehicles are connected if their distance is smaller than r. Let L_c be the gap between two connected neighboring vehicles. We have

$$f_{L_c}(l) = f_L(l \mid L < r) = \begin{cases} \dfrac{f_L(l)}{F_L(r)} & r \geq l > 0 \\ 0 & \text{else} \end{cases} \tag{2-31}$$

Let M be the number of vehicle gaps in a partition,

$$Y = \sum_{i=1}^{M} L_{c_i} \tag{2-32}$$

Taking the average of both sides of the above equation, we get

$$E[Y] = E[M] * E[L_c] \tag{2-33}$$

while M has a geometric distribution

$$P[M = n] = F_L(r)^n * (1 - F_L(r)) \tag{2-34}$$

With the above derivation of $E[T_c]$, $E[X_c]$, $E[T_f]$ and $E[X_f]$, we can compute v_p using equation (2-13). One remaining problem is the vehicle gap distribution. It is shown by Rucack et al. [70] that the gap distribution can be approximated by an exponential distribution with the vehicle density $\mu = \lambda / E[V]$ where μ represents the number of vehicles per unit of distance and λ is vehicle traffic flow rate. The positions of vehicles are distributed according to a Poisson point process as a result.

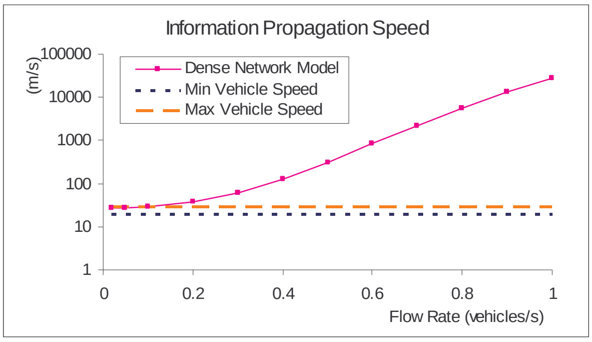
2.2.6 Between "Sparse" and "Dense"

So far we have presented models for sparse networks and dense networks. For networks between these two extremes, the sparse network model is not appropriate since the message propagation within partitions is ignored. Neither is the dense network model appropriate because the assumption that vehicle traffic between cycles is independent does not hold. The sparse network model provides a lower bound on propagation speed because the fast propagation within partitions is ignored. The relationship between the value predicted by the dense network model and the actual value is not immediately apparent. Let us examine Figure 4(c) again. During the next catch-up process, by assuming a uniform stream of vehicles passing through location H', there are two opposing factors. First, the average speed of vehicles close to location C in set 1 is underestimated, which tends to lower the computed information propagation speed. Second, the average speed of vehicles close to location C from set 2 is augmented, which inflates the computation result. As H' is far away from C, the influence of both factors diminishes and the dense network model result should approach the expected result.

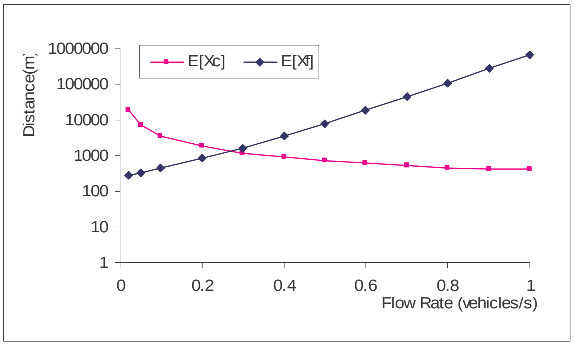
2.2.7 Numerical Results

These models provide us with the ability to understand and explore information propagation under different traffic conditions. In this section, we use these models to examine the sensitivity of various traffic parameters. It can be seen from the above analysis that the catch-up time T_c is determined by the relative speed between vehicles, the vehicle flow rate, and the vehicle density. The message propagation distance X_c during the catch-up phase is determined by both the catch-up time and the absolute vehicle speed. The partition size Y and thus the message propagation distance during the forward phase X_f are determined by vehicle density. The time spent in the forward phase T_f can be neglected given the relative speed of message propagation vs. vehicle movement. Next, we derive numerical results using the above models to verify these claims. We assume vehicle speed distribution is V ~ uniform(v_{min}, v_{max}), where v_{min} = 20m/s (45mph), v_{max} =

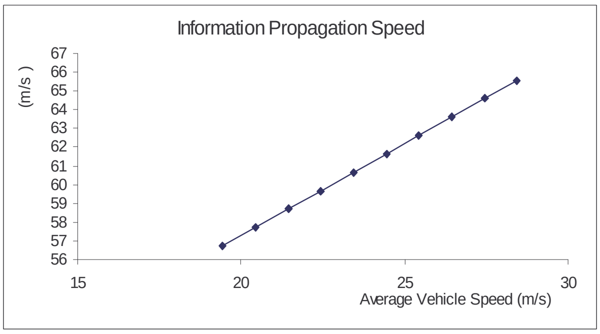
29m/s (65mph), the traffic flow rate λ = 1080vph, the radio range r = 250m, and the message processing time tr = 4ms as a starting point and vary individual parameters later.



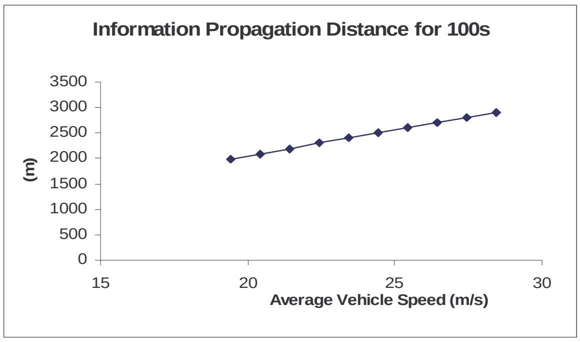
(a)



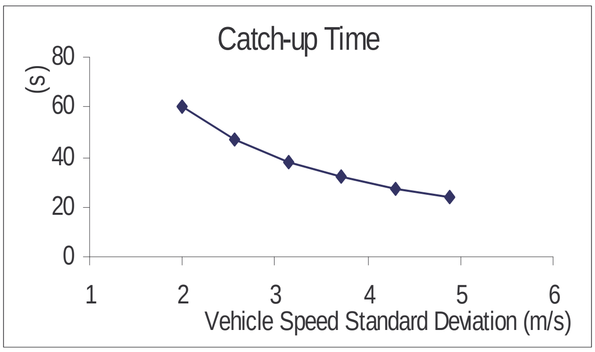
(b)



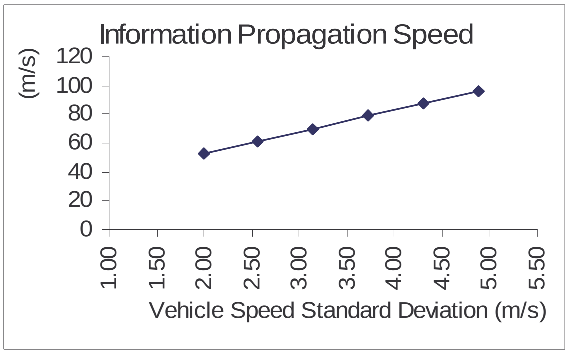
(c)



(d)



(e)



(f)

Figure 8: Numerical results

We start by showing how the traffic density impacts the propagation (equivalent to varying the vehicle radio range) while fixing other parameters. Since the vehicle speed distribution remains the same, the vehicle density is proportional to the vehicle flow rate. We thus vary the vehicle flow rate. Figure 8(a) plots the information propagation speed as predicted by the dense network model as the function of vehicle flow rate. Because higher vehicle density leads to a larger partition size, shorter inter-partition distance and thus shorter catch-up time, it is not surprising that the propagation speed increases sharply as the vehicle flow rate increases and can be much faster than vehicle movement. Although it is not easy to see from the figure, while the vehicle flow rate is very low (<0.05 vehicles/s) the predicted information propagation speed is less than the maximum vehicle speed. Based on the analysis in Chapter 2.2.4, the information propagation speed predicted by the sparse network model can approach the maximum vehicle speed over time. This suggests that when the network is very sparse, the sparse network model is a better estimator over a long time though it still provides a lower bound. The opposite is true otherwise. In Figure 8(b), we compare $E[X_c]$ and $E[X_f]$. If the source and destination of a message are both within the same partition, there is an end-to-end (E2E) connection. Many applications require high E2E connectivity. As expected, the probability of E2E connectivity increases as the vehicle density increases (as indicated by the increasing $E[X_f]$ and decreasing $E[X_c]$). Once the vehicle flow rate reaches 0.6 vehicles/s (where $E[X_f] >> E[X_c]$), E2E connectivity exists with high probability.

We next show the impact of the mean vehicle speed. We vary the mean vehicle speed while maintaining the speed variance among vehicles and the vehicle density μ (which also implies the vehicle flow rate λ is varied accordingly since $\mu = \lambda / E[V]$). We first set $\lambda > 0.2$ vehicles/s and study the dense network model. Just as expected, the mean catch-up time remains unchanged at 46.7 seconds. As the mean vehicle speed increases the propagation distance during the catch-up phase increases. The average message propagation speed increases somewhat as a result (see Figure 8(c)). We then set $\lambda = 0.05$ vehicles/s and study the sparse network model. As shown in Figure 8(d), the propagation distance for 100 seconds increases as the result of the increase of the mean vehicle speed because message propagation is assumed to rely solely on vehicle movement in the sparse network model.

We next examine the effect of the relative speed between vehicles. We vary the speed variance among vehicles but keep the average vehicle speed constant. We set $\lambda = 0.3$ vehicle/s and only the dense network model is explored. Figure 8(e) and (f) plot the mean catch-up time and information propagation speed, respectively. The results confirm our claim that larger relative speed between vehicles leads to quicker catch-ups, and thus higher message propagation speed overall. These results also match the observation by Chen et al. in [17].

2.3 Models for Two-way Vehicle Traffic

In this section, we consider the models for two-way vehicle traffic. They are extensions of those for one-way vehicle traffic described earlier.

2.3.1 Two-way Vehicle Traffic

There are two traffic flows in opposite directions, called the positive flow and negative flow, respectively. Vehicles in one direction can help with message propagation in the opposite direction. For example, as shown in Figure 9, vehicle A cannot directly talk to vehicle B. But the message can propagate to vehicle B through vehicle C traveling in the opposite direction. The two-way vehicle traffic facilitates message propagation in two ways: first, it extends the partition and shortens the inter-partition distance by increasing the vehicle density; second, it speeds up the catch-up process by increasing relative vehicle movement.

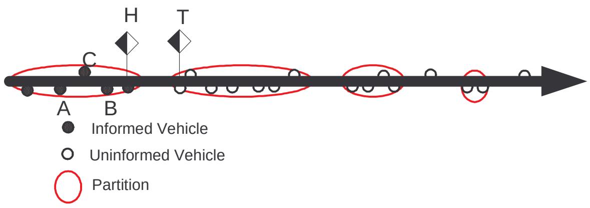


Figure 9: Beginning of a catch-up process

We use the subscript + (-) to represent vehicle traffic parameters in the positive (negative) direction, e.g., V_+ (V_-) is vehicle speed in the positive (negative) direction. Suppose $V_+>0$ and $V_-<0$, and

$$V_+ \in [v_{min+}, v_{max+}], \; V_- \in [-v_{max-}, -v_{min-}]$$
$$\text{where } v_{min+}, v_{max+}, v_{min-}, v_{max-} > 0$$

Without loss of generality, we examine message propagation in the positive direction.

2.3.2 Sparse Network Model

With the sparse network model, message propagation is assumed to rely solely on vehicle movement so that the negative traffic flow does not help propagate the message in the positive direction. The model in Chapter 2.2.4 is still a good approximation when the network is sparse in both directions[1].

2.3.3 Dense Network Model

With the assumption of Poisson processes in both directions, we can model vehicles passing an arbitrary point on the road as two independent Poisson processes with rate λ_+ and λ_-, respectively. Similarly the positions of vehicles at any instant are distributed according to two independent Poisson point processes with vehicle density $\mu_+ = \lambda_+/E[V_+]$ and $\mu_- = \lambda_-/(-E[V_-])$, respectively. Compared with the one-way traffic scenario, the gap distribution, the distribution of the message head traveling a distance X(t) and the partition tail traveling a distance X'(t) during the catch-up phase need to be re-examined. Other procedures are the same as in the one-way traffic model.

Taking vehicles in both directions into account, the inter-vehicle gap has an exponential distribution with the rate $(\mu_+ + \mu_-)$.

We now explore the distribution of X(t). The system configuration at the beginning of a cycle is plotted in Figure 9. Let *ve* be the vehicle at location H at time 0, whose average speed is V_0, S_+ denotes the set of vehicles passing location H in the positive direction during (0,t] with cardinality $N_+(t)$, and S_- denotes all vehicles passing location H in the negative direction during the time interval $[-(v_{max-}/\ v_{min-}-1)*t,\ 0)$ with cardinality $N_-(t)$. If $N_+(t) = 0$, the message head at time t can be approximated as the vehicle of the maximum position among *ve* and all vehicles in S_-. If $N_+(t) > 0$, the message head at time t is the one of the maximum position among *ve* and all vehicles in S_+. This yields

$$X(t)=\begin{cases}\max(V_0*t,\ \max(V_{j-}*(t-T_{j-})))\ \ j=1,2,...,N_-(t) & N_+(t)=0\\ \max(V_0*t,\ \max(V_{i+}*(t-T_{i+})))\ \ i=1,2,...,N_+(t) & N_+(t)>0\\ 0 \quad \text{else}\end{cases} \qquad (2\text{-}35)$$

while $T_{i+}(T_{j-})$ is the time when vehicle i (j) passes location H in the positive (negative) direction.

Conditioned on $N_+(t)$ and $N_-(t)$, we have

[1] Sometimes we have high vehicle density in one direction and low in another direction. The dense network model is more appropriate for this type of situations.

$$F_{X(t)}(x) = P[N_+(t)=0] * \sum_{n=0}^{\infty} P[X(t)<x \mid N_+(t)=0,\ N_-(t)=n] * P[N_-(t)=n]$$
$$+ \sum_{n=1}^{\infty} P[X(t)<x \mid N_+(t)=n] * P[N_+(t)=n] \tag{2-36}$$

Remember that two flows pass location H as two independent Poisson process and vehicles are independent. We have

$$\begin{aligned} & P[X(t) < x \mid N_+(t) = 0, N_-(t) = n] \\ &= P[V_0 * t < x,\ V_{j-} * (t - T_{j-}) < x \text{ for each } j = 1, 2, ..., n] \\ &= P[V * t < x] * P[V_- * (t - T_-) < x]^n \end{aligned}$$

$$\begin{aligned} & P[X(t) < x \mid N_+(t) = n] \qquad n > 0 \\ &= P[V_0 * t < x, V_{i+} * (t - T_{i+}) < x \text{ for each } i = 1, 2, ..., n] \\ &= P[V * t < x] * P[V_+ * (t - T_+) < x]^n \end{aligned} \tag{2-37}$$

While with properties of Poisson process, we have

$$T_+ \sim \text{uniform}(0,\ t), \quad T_- \sim \text{uniform}(-(v_{max-} / v_{min-} - 1) * t, 0) \tag{2-38}$$

$$P[N_-(t) = n] = \frac{e^{-\lambda_- * (v_{max-}/v_{min-} - 1) * t} * (\lambda_- * (v_{max-} / v_{min-} - 1) * t)^n}{n!}$$

$$P[N_+(t) = n] = \frac{e^{-\lambda_+ * t} * (\lambda_+ * t)^n}{n!} \tag{2-39}$$

Simplifying equation (2-36), we get

$$F_{X(t)}(x) = \begin{cases} P[N_+(t)=0] * P[V*t<x] * \sum_{n=0}^{\infty} P[V_- * (t-T_-)<x]^n * P[N_-(t)=n] & x<0 \\ P[V*t<x] * \sum_{n=0}^{\infty} P[V_+ * (t-T_+)<x]^n * P[N_+(t)=n] & \text{else} \end{cases} \tag{2-40}$$

The distribution of speed V of a random vehicle on the road can be represented as

$$f_V(v) = \frac{\mu_+}{\mu_+ + \mu_-} * f_{V+}(v) + \frac{\mu_-}{\mu_+ + \mu_-} * f_{V-}(v) \tag{2-41}$$

Thus given the distribution of V_+ and V_-, we can derive $F_{X(t)}(t)$ as in equation (2-40). Similar analysis can also be done for the distribution of X'(t). The final result is shown in equation (2-42).

$$P[X'(t)>x]=\begin{cases} P[N_-(t)=0]*P[V*t>x]*\sum_{n=0}^{\infty}P[V_+*(t-T_+)>x]^n*P[N_+(t)=n] & x>0 \\ P[V*t>x]*\sum_{n=0}^{\infty}P[V_-*(t-T_-)>x]^n*P[N_-(t)=n] & \text{else} \end{cases} \quad (2\text{-}42)$$

$T_- \sim \text{uniform}(0,\ t),\quad T_+ \sim \text{uniform}(-(v_{max-}/\ v_{min-}\ -1)*t, 0)$

Other analysis procedures follow Chapter 2.2.5. Eventually we can compute v_p, the long-term average message propagation speed as equation (2-13).

2.3.4 Verification

We use simulation to verify the preceding two-way dense network model. We use the distributed simulation test bed for vehicular network analysis mentioned in Chapter 1.4. We model a two-way freeway with four lanes in each direction. Vehicles enter the system at the rate of 5000vph on each way and the free flow speed is 50mph. We implement the idealized data propagation scheme described in Chapter 2.2.1, i.e., messages are greedily propagated among vehicles. Protocol layers and signal interference are not modeled. A message begins to propagate from a randomly selected vehicle on the road. The road is populated with vehicles before the message begins to propagate. Vehicle locations are updated in each time step. The vehicle radio range r = 250m, and the message processing time tr = 4ms. We measure the message propagation distance of 10km in one direction. We believe an information propagation distance of 10km is reasonable for many ITS applications (e.g., vehicle traffic information). We evaluate different vehicle traffic densities by varying the penetration ratio. The dense network model requires the vehicle density to be sufficiently dense, thus limiting the lower bound of the vehicle density we can explore (we set the upper bound of the inter-vehicle gap as the radio range based on previous analysis). The limited propagation distance chosen (10km) determines the upper bound of the vehicle density (big number theory) we can explore. Before the simulation we use analytical models to derive useful parameter settings. Each reported result below is the average of 40 simulation measurements with different random number seeds.

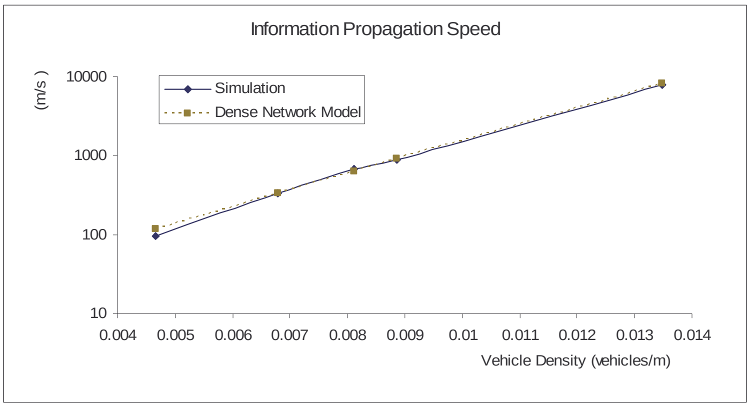


Figure 10: Simulation verification

In Figure 10, we plot the information propagation speed for both the analytical model and the simulation as the function of vehicle density. It shows that the analysis and simulation results are reasonably close to each other, suggesting the assumptions we made for this model are reasonable.

2.3.5 Discussion

In Chapter 2.2, we presented models for one-way vehicle traffic where information is propagating in the direction of vehicle traffic flow. The models for two-way vehicle traffic allow one to explore the scenario that information is traveling in the opposite direction of the vehicle traffic by assuming there is no positive vehicle traffic flow.

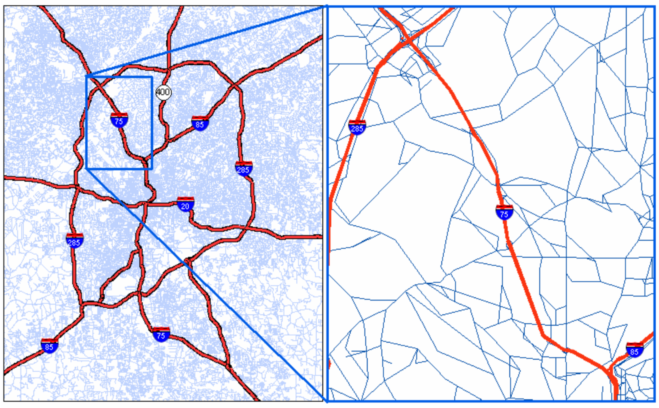
Information is able to propagate with only an opposite vehicle traffic flow as long as the backward traveling distance of the message head during the catch-up phase is less than the forward traveling distance during the forward phase on average.

2.4 Simulation Study

In this section, we describe simulation results for multi-hop message propagation in a more complex highway scenario.

2.4.1 Experimental Design

The simulation infrastructure used to perform this study has been introduced in Chapter 1.4. The modeled area for this research effort is the I-75 corridor in the northwest quadrant of Atlanta, Georgia, traversing I-75 from the I-85 interchange to the south to the I-285 interchange to the north (Figure 11). The modeled area incorporates approximately 7.6 miles of I-75 with a posted speed limit of 55 mph. In addition to the freeway, approximately 100 miles of arterials surface streets are included within the study area.



(a) (b)

Figure 11: **(a) I-75 corridor study area in greater Atlanta region (b) Corridor study area**

The spatial propagation of information southbound along I-75 for a distance of 6 miles is simulated in this research. Vehicle traffic in both directions is exploited in relaying the message. Again, we employ an idealized data propagation scheme as described in Chapter 2.2.1: communications range r is set as 250m (within the typical clear path range of an 802.11 communication system); the time for receiving and processing a message tr is set at 4ms. Message propagation was simulated under two traffic scenarios (evening peak and nighttime traffic, with typical traffic volumes derived from the regional travel demand model) and for various fleet penetration ratios. CORSIM was utilized to generate 600-second traffic traces, using a different random number seed for each trace. The simulation results for all scenarios (except for penetration ratio of 0.05), are obtained by partitioning each 600-second trace into five 100-second samples. Each sample then begins with the propagation of a single message. For each penetration ratio, 10 traces are examined for a total of 50 samples. For the 0.05 fleet penetration ratio, 20 traces were generated and one sample was collected from each trace. This arrangement of random number seed variation and 100 second samples was utilized to lessen the correlation between successive message propagations.

2.4.2 Results

2.4.2.1 Average End-to-End Delay

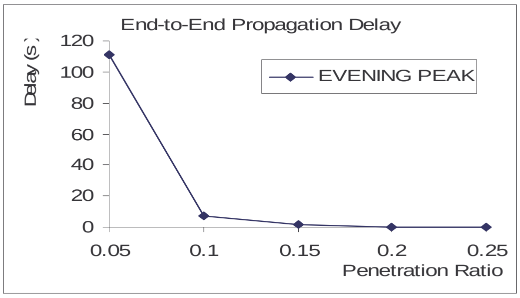


Figure 12: Evening peak end-to-end propagation delay

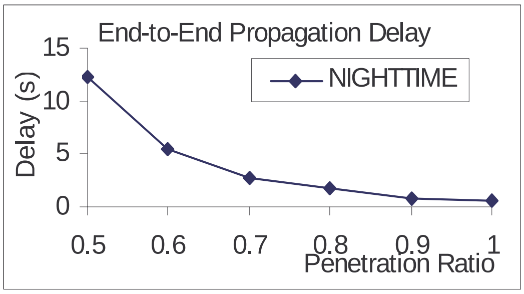


Figure 13: Nighttime end-to-end propagation delay

Figure 12 and Figure 13 illustrate the average E2E propagation delay as a function of the fleet penetration ratio for evening peak and nighttime traffic, respectively. Note that for evening peak traffic, the penetration ratio shown starts from 0.05 while for nighttime traffic, it starts from 0.5. This is because when the penetration ratio is below 0.5 for nighttime traffic, some samples require a propagation time of much more than 100 seconds. One can expect the delay is much higher when the penetration ratio is below 0.5. When the fleet penetration ratio is below 0.10, our trials show that information propagation is principally driven by vehicle movement. This agrees with the analysis in Chapter 2.2.2. Overall the delay tends to decrease as the fleet penetration ratio increases, but not linearly. To achieve an average delay below 2 seconds for the 6-mile message propagation, a fleet penetration ratio of 0.15 or greater is required for evening peak traffic, while the penetration ratio has to reach 0.80 for nighttime traffic (given the sparse traffic volumes on this freeway at night). Since intra-partition vehicle forwarding is relatively fast, the majority of the propagation time is spent moving information from one partition to another. The shortest E2E delay for 6 miles observed is 0.164 seconds given the parameters used here, where an E2E path exists. During evening peak traffic (Figure 12), message delay reaches a minimum at a fleet penetration ratio of around 0.20. Further fleet penetration beyond 0.20 does not significantly reduce propagation delay.

2.4.2.2 Delay Variance

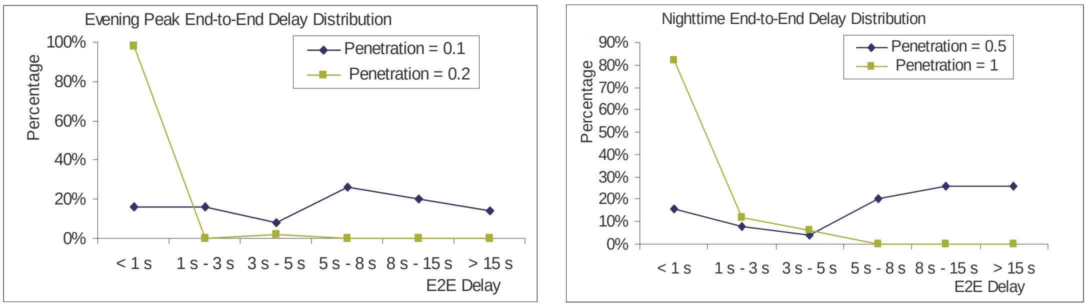


Figure 14: End-to-end delay distribution

Figure 14 shows the message propagation time distribution for several penetration ratios. As expected, the propagation time varies significantly. This variation is primarily due to path vulnerability, i.e. the likelihood of partitioning. When a contiguous E2E path exists, message propagation can finish within one second. When an E2E path does not exist, (i.e., the communication network contains separated partitions) seconds or even minutes may be necessary for a message to propagate through the chosen distance, depending on the number of partitions traversed and the distance between partitions. Only with a high density of instrumented vehicles where the E2E paths almost always exist (e.g., penetration ratio > 0.20 for evening peak traffic) do we observe stable performance. Thus, a message only has a higher probability of rapid propagation in an environment of higher density of instrumented vehicles.

2.4.2.3 E2E Connectivity

Most routing protocols for ad hoc networks assume E2E connectivity [38, 54, 79]. One natural question is then: under what circumstances can E2E connectivity be safely assumed in V2V networks? Here we show some evaluation results of E2E connectivity.

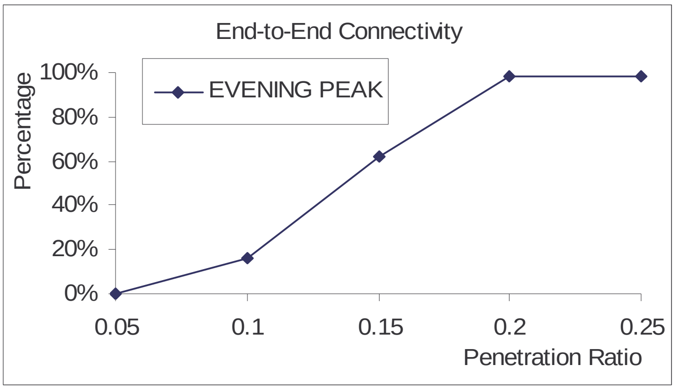


Figure 15: Evening peak end-to-end connectivity

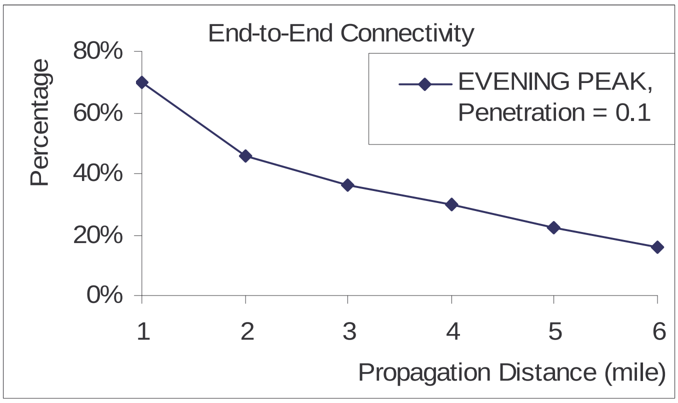


Figure 16: Evening peak end-to-end connectivity for fleet penetration ratio of 0.1

Figure 15 shows the percentage of propagation paths where there is an E2E path for evening peak traffic. When the fleet penetration ratio exceeds 0.20, the connectivity can reach nearly 100%. When the penetration ratio is lower than 0.10, the connectivity is below 15%. However the connectivity is only 82% even with 100% fleet penetration (i.e. every vehicle is instrumented) for nighttime traffic. Traffic volumes are so low in both directions that distance separations create multiple partitions. Figure 16 illustrates the E2E connectivity versus the propagation distance when the fleet penetration ratio is 0.10 for evening peak traffic. Dousse et al. [21] demonstrated that E2E connectivity decreases with distance for one-dimensional network topologies. Not surprisingly, our results also reflect this trend. When the propagation distance is 1 mile, the E2E connectivity can reach about 70%, while it is only 16% for the propagation distance of 6 miles.

2.4.3 Discussion of Results

Several observations may be made based on the results above.

First, V2V communications appear feasible for propagating certain types of information (e.g., traffic and traveler information) along I-75 freeway in the Atlanta metro areas, as well as other roadway systems with similar traffic characteristics as Atlanta during peak or high traffic density periods (times during which the propagation of traffic information is most needed). The propagation performance depends largely on the density of instrumented vehicles along the E2E path, which is a function of the traffic density and fleet penetration ratio. With a sufficient fleet penetration ratio and traffic flow rate (penetration ratio greater than 0.2 during daytime), information can quickly propagate through the system. When the penetration is high, even some interactive applications (e.g., games and chat) might be supported between vehicles over multi hops. Rapid message propagation during low traffic density periods, e.g., nighttime, presents challenges and may require some additional mechanism to support communications, e.g., deployment of fixed-location roadside relay stations.

Second, the simulation methodology described here allows one to determine a target penetration ratio for effective communications as a function of application requirements and traffic density. For example, when rapid message propagation is desired during the evening peak, a penetration ratio of approximately 0.20 is sufficient for effective information propagation.

Third, the message propagation delay is highly variable except when vehicle density becomes saturated. A particular delay may be well below or above the average, depending on prevailing traffic conditions. For applications requiring highly reliable, minimal message propagation times it may be necessary to design networks that provide extra support to avoid such variations. For example, to reduce path vulnerability roadside relays could supplement the communication infrastructure in critical areas. Or, a subset of vehicles could be equipped with cellular messaging systems, through which critical information would be reliably relayed without being dependent upon a vehicle-to-vehicle communication. For applications where immediate data are not critical, other non-fixed-infrastructure based solutions may be explored. For example, vehicles can cache data and use this information when up-to-date information is not available.

Lastly, E2E distance is an important factor. E2E connectivity is possible over long distances when instrumented vehicle density is high, or at slightly lower densities when propagation distances are short.

2.4.4 Lane Closing Message Propagation

Another evaluation conerns the propagation of information in a highway crash scenario. The scenario involves an accident in the northbound lanes of I-75 near the northern end of the modeled region during the evening rush hour. This accident results in the closure of two lanes (the second and third from the left). An informative message is generated by one of the vehicles involved in the crash, and the information propagated southward along I-75 using V2V communications. The message propagates a distance of 6 miles. The results below are the aggregation of 50 samples with different random number seeds.

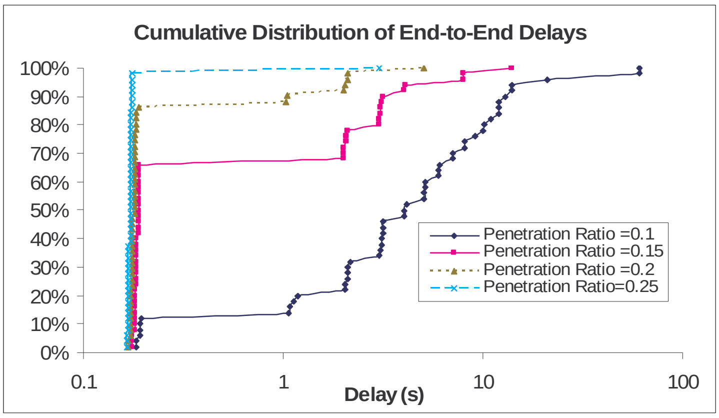


Figure 17: Cumulative distribution of end-to-end delays

The average E2E propagation delays for penetration ratios of 0.1, 0.15, 0.2 and 0.25 are 7.62624, 1.49832, 0.36216, and 0.1708 seconds, respectively. The shortest E2E delay for the 6 mile propagation distance was observed to be 0.164 seconds given the parameters used here where an E2E path exists. When the penetration ratio reaches 0.25, the average E2E delay is very close to this minimum and the message can quickly propagate through the designated area. Further penetration will not result in much faster propagation. When the penetration ratio is low (<0.2), the majority of the propagation time is spent moving information from one partition to another since intra-partition vehicle forwarding is relatively fast. Figure 17 shows the cumulative distribution of E2E delays. When the penetration ratio is 0.1 or 0.15, a large variance in delay was observed. This is the result of a high probability of network partitioning. Once the penetration ratio reaches 0.2, an E2E path usually exists and thus a rapid propagation occurs most of the time. When the penetration ratio is 0.25, it is very likely that an E2E path exists.

The above results show the feasibility of using V2V communications to propagate information along the I-75 freeway in the Atlanta metro areas, as well as other roadway systems with similar traffic characteristics as Atlanta during peak or high traffic density periods, even with a relatively low penetration ratio (0.2).

2.5 Conclusion

In this chapter, we develop a collection of models to study information propagation using V2V communications. These models can provide expected information propagation speed/distance over time. We also provide an initial investigation needed to implement and test the feasibility of V2V communication networks through simulation. Simulation study allows the exploration of a much richer set of performance metrics than analytical models, including delay distribution (out of which average, variance, and others quantities of interest can be derived) and E2E connectivity. In prior work, information propagation through either multi-hop forwarding or mobile node movement was examined. We take into consideration that information can propagate in both modes.

Our analytical models capture the basic dynamics of the vehicle traffic – vehicles move along roads and slower vehicles are overtaken by faster vehicles. For the sparse network, information propagation is modeled as solely depending on vehicle movement. For the dense network, information propagation is modeled as a renewal reward process. We cover both one-way vehicle traffic and two-way vehicle traffic scenarios. Our models reveal several important vehicle traffic parameters significantly impacting information propagation: the vehicle density, average vehicle speed, and relative speed among vehicles. The vehicle density determines partition size and inter-partition distance. The average speed decides the message propagation speed during the catch-up process. The relative speed among vehicles affects the time to catch up another partition.

Several research problems could benefit from these models:

1. System architecture design. In a hybrid network architecture, wireless base stations are placed in strategic locations. Vehicles can communicate with both base stations and peers. Vehicles out of the range of a base station can only communicate with it through the relay of other vehicles and/or other base stations. Our models can be used to help design a placement strategy where the average delay to reach a vehicle at any location is bounded by some prescribed threshold.
2. Data dissemination. When a message needs to be delivered from one location to another, the path with the shortest transmission delay is often desired. Our models can help the path selection strategy in estimating the information propagation speed along different paths. When disseminating data in V2V networks, flooding is expected to be performed frequently due to the high vehicle mobility. Our models can be employed to estimate the outreach of messages so that flooding in places far behind can be quenched.

CHAPTER

3 DATA DISSEMINATION ALGORITHMS

3.1 Motivation

Deployment of applications on various vehicular networks requires the support of data dissemination services. Data dissemination concerns the transport of information to *intended receivers* while meeting certain *design objectives*, e.g., delay, loss rate, overhead, etc. In this chapter, we study how instrumented vehicles can exploit short-range communications (e.g., DSRC or 802.11x) between vehicles to disseminate information. Utilizing V2V communications offers the merits of low cost and easy deployment. Both the "pure ad-hoc" and "hybrid" architectures mentioned in Chapter 1.1 require the use of V2V communications to realize data dissemination. Multi-hop forwarding is often considered a low-cost alternative to the more expensive infrastructure-based approach.

Based on the analysis in the preceding chapter, vehicular networks relying on short-range communications are usually partitioned (or at least not continuously connected). Store-carry-forward message switching allows nodes to store messages temporarily and forward them later. In Delay-Tolerant Network research [23], store-carry-forward message switching is proposed to overcome the problems associated with intermittent connectivity, long or variable delay, and high error rates. **Opportunistic forwarding** is one type of store-carry-forward scheme where data forwarding happens during opportunistic contacts (as opposed to scheduled contacts). Since the future vehicle contacts are usually unpredictable, opportunistic forwarding, by itself, is a viable approach to disseminate information using V2V communications for applications that can tolerate some data loss and delay.

In this chapter, we first present a generic design methodology for enhanced opportunistic forwarding algorithms. We then derive two such algorithms, MDDV (Mobility-centric Data Dissemination algorithm for Vehicular networks)[87] and *optimistic forwarding* [17], from this methodology. We evaluate their performance using a detailed vehicle traffic model of a portion of the city of Atlanta. We discuss algorithm design largely in the context of ad hoc V2V networks, but our design is also applicable in hybrid architectures when addressing data dissemination beyond the fixed infrastructure.

3.2 Data Dissemination Services

Data dissemination concerns the transport of information to *intended receivers* while meeting certain *design objectives*. The design objectives considered here include low delay, high reliability, and low message overhead. Message overhead is defined as the number of message transmissions. The intended receivers are those specifying interest in the information. The intended receivers may be defined arbitrarily: "all vehicles going to the football stadium", "emergency vehicles that are close by", etc. Here, we are only concerned with those definitions that can be readily exploited by data dissemination algorithms, i.e. those dependent on time and location.

An important question concerns the semantics of data dissemination services, and their suitability for possible applications. Four services that have immediate application are unicast, multicast, anycast and scan. (1) Unicast with *precise* location means a message should be delivered to vehicle *i* in location *l* before time *t* (t is the expiration time of the message which can be used to prevent indefinite message looping in a system.) Unicast with *approximate* location means sending a message to vehicle *i* before time *t* while that vehicle was last known to be at location *l* with mobility *m* at time *t'*, *where t'<t*. (2) Multicast means disseminating a message to all reachable vehicles in region *r* between the time when the message arrives at the region and time *t*. (3) Anycast means disseminating a message to one among a set of possible receivers (e.g., send to any emergency vehicle) in region *r* before time *t*. (4) Scan is to have a message traverse region *r* once before time *t*. In these services, location *l* and region *r* are used to direct the message to a geographical area. Time *t* (which implicitly determines message life span) is determined by the nature of the message, e.g., when the information becomes obsolete. Other services can also be designed as variations or combinations of the above services.

To illustrate an application using these services, consider a vehicle (or a traffic signal controller) wishing to obtain vehicle traffic information concerning some remote region. The vehicle/controller desiring the information first queries its own proximity (*multicast*) to determine if a near-by vehicle happens to have this information. Any vehicle having such information can respond (*unicast with approximate/precise location*). If no vehicle replies within a certain amount of time, the vehicle/controller sends a query to any vehicle in that particular remote region (*anycast*). Any receiver in the remote region can respond. The response can be disseminated as *unicast with approximate/precise location*, or *multicast* if caching is desired. This scenario describes a pull approach. A push approach scenario may also be envisioned, e.g., vehicles encountering a crash or traffic congestion may send this information to an upstream or downstream region using *multicast*.

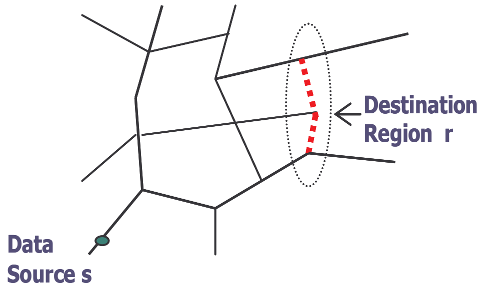


Figure 18: Geographical-temporal multicast

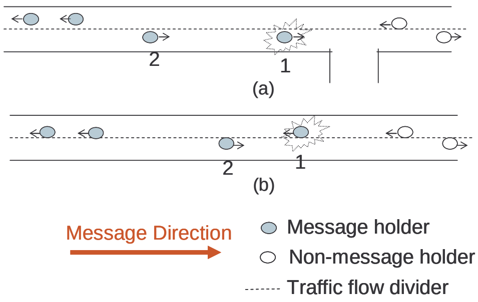


Figure 19: Lack of perfect knowledge

In this chapter, we illustrate our design by reference to a test case - **geographical-temporal multicast**. As will be shown later all the services mentioned above can be easily covered with extensions of the design for this test case. Geographical-temporal multicast is formally defined as: *deliver a message to all reachable vehicles in region r between the time when the message arrives at the region and time t, with an additional constraint that the data source s is outside of r*. A region is defined as a set of connected road segments (for two-way roads, both directions are included). Two road segments are connected if they share an intersection (see Figure 18 for an example). The source is placed outside of the destination region to make the problem non-trivial.

3.3 Enhanced Opportunistic Forwarding Design

The enhanced opportunistic forwarding algorithms discussed in this chapter combine simple opportunistic, geographical, and trajectory based forwarding (see Chapter 1.5.1.4). A directed forwarding trajectory is specified extending from the source point to the destination region (*trajectory base forwarding*) along the road graph, over which a message will be moved geographically closer to the destination region (*geographical* *forwarding*) through vehicle movement or data forwarding during opportunistic contacts *(opportunistic forwarding)*. In this chapter, we illustrate the design through addressing disseminating a single message.

3.3.1 Assumptions

Given the performance of current onboard telematics technologies, we assume a vehicle has knowledge of the road topology through a digital map and its own location in the road network via a GPS device. We also assume vehicles have knowledge of the existence of their neighbors through some network link level mechanism. Our algorithms do not require vehicles to know the exact location of neighbors, but can exploit the information for improved performance and functionality when it is available. Most geographical forwarding algorithms assume the knowledge of exact neighbor location. However, a vehicle's impression of neighbor location might be inaccurate either due to high vehicle mobility or absence of neighbor location information. For simplicity of discussion it is assumed that all vehicles communicate using the same wireless channel[2] (e.g., one of the service channels in DSRC). The message dissemination information (e.g., source id, source location, generation time, destination region, expiration time, and forwarding trajectory) is specified by the data source and is placed in the message header.

3.3.2 A Generic Design Methodology

As discussed in Chapter 1.2 it is desirable that applications running on V2V networks be designed using localized algorithms where nodes perform local operations and interact with neighbors while their collective behavior achieves some global objective. This approach addresses the inherent uncertainties of V2V networks and scales well as the system size increases. Our generic methodology for designing localized algorithms in V2V networks consists of four steps: specification of a desired global behavior, specification of an ideal scenario, definition of approximations to deal with incomplete and/or inaccurate information, and design of local operations to implement approximations. The global behavior is the objective desired. The ideal scenario characterizes algorithm behavior by assuming every participant has perfect knowledge of the entire system. An approximation is a practical scheme used to approach the ideal scenario where the perfect knowledge assumption is relaxed. Local operations are designed to realize the approximation.

Global Behavior. The geographical-temporal multicast utilizes data dissemination along a specified directed forwarding trajectory. The directed forwarding trajectory is specified by the message source. For example, in this research the forwarding trajectory is specified as the shortest road path that a vehicle may take to reach the destination region. Future research will explore more complicated trajectory creation strategies, such as those accounting for the road topologies and relative dispersion of instrumented vehicles. The data dissemination process consists of two phases: the forwarding phase and the propagation phase. In the forwarding phase, the message is forwarded or carried by message holders along the forwarding trajectory to the destination region. Because message forwarding is significantly faster than vehicle movement, it is desired to exploit forwarding opportunities in order to reduce message delay. Once the message reaches the destination region, the propagation phase begins and the message is propagated to every vehicle in the destination region until the message lifetime expires.

---

[2] This condition is not necessary as multiple short-range technologies can be exploited in data dissemination

Ideal Scenario. During the forwarding phase, the message holder closest to the destination region along the forwarding trajectory is called the "message head." The vehicle taking the role of message head may change over time as a result of message transmissions or vehicle movement. Under perfect knowledge conditions, every vehicle may identify the message head in real time. Given this perfect knowledge the ideal algorithm may be designed such that only the message head actively seeks to pass the message to other vehicles that may be closer to the destination region (possibly changing the message head vehicle). During the propagation phase, the message is propagated to vehicles that have not received the message in the specified destination region.

Approximation. The above ideal scenario is not a realistic reflection of real-world conditions as vehicles do not have perfect knowledge of the entire system. Thus individual vehicles cannot reliably identify the message head vehicle in real time, even when they are the message head themselves. For example, as illustrated in Figure 19, in a road carrying two-way traffic, the current message head is vehicle 1. In (a), vehicle 1 may travel away from the forwarding trajectory or may become inoperative. When this happens, vehicle 2, the immediate follower, should become the new message head but may not be aware of this fact due to network partitioning. In (b), vehicle 1 is the current message head. When vehicle 2 passes vehicle 1, vehicle 2 should become the new message head. However, vehicle 2 does not know this unless it receives an explicit notification from vehicle 1. Should no vehicles claim to be the message head, the message must rely on vehicle movement to reach the destination region, potentially resulting in long delays.

simultaneously.

One naïve approach to avoid the loss of the message head is to have every message holder actively seek to forward the message. The drawback of this approach is the potential for significant message overhead. In an attempt to reduce excessive message overhead, while maintaining reliable message delivery, we have designed two different algorithms, each intended to approximate the ideal scenario. These algorithms, MDDV and *optimistic forwarding*, differ during the forwarding phase but are the same during the propagation phase.

MDDV employs the concept of "*group forwarding.*" A group of vehicles near the real message head are identified to actively forward the message. Multiple vehicles actively forwarding the message help to mitigate the uncertainties inherent in V2V networks. Limiting the number of vehicles actively forwarding the message to a small group near the real message head ensures that the overhead will not be too high. MDDV allows the group membership to change as the actual message head moves toward the destination region, as will be explained later. MDDV addresses both scenarios in Figure 19.

Optimistic forwarding is designed to closely approximate the ideal scenario. Each message is assigned an owner during the forwarding phase. Efforts are made to try to make the message head the owner. Only the message owner may transmit the message. Unlike the ideal scenario, the owner cannot be identified automatically but has to be transferred explicitly through communications. *Optimistic forwarding* can deal with the scenario in Figure 19(b) but fails to address the scenario in Figure 19(a) if the network is partitioned or the message owner fails.

Intuitively MDDV can perform better than *optimistic forwarding* but may incur higher message overhead due to message replication. Next we will first describe MDDV and *optimistic forwarding* during the forwarding phase. We then brief discuss their common behavior during the propagation phase.

3.4 MDDV

MDDV approximates the ideal scenario by identifying a group of vehicles near the real message head to actively forward the message. Due to the distributed nature of the algorithm, each vehicle decides whether it belongs to this group based on its own knowledge. Vehicles utilize two pieces of information in making this group membership determination: localized vehicle traffic conditions and an approximation of the message head location. For example, if the instrumented vehicle density is high, vehicles actively forwarding the message should be in close proximity to the real message head, or the group size will be larger than is necessary, resulting in excess message overhead. In this section, we will describe how vehicles obtain and apply the necessary information.

3.4.1 Vehicle Traffic Information

As explained in Chapter 1.5.3, the three fundamental characteristics of vehicle traffic are flow q (vehicles/hour), speed u (km/h) and density λ (vehicles/km). The average values of these quantities can be approximately related by the basic traffic stream model $u=q / \lambda$. The actual relationships are much more complex, and are usually modeled in simulation models.

Vehicles can monitor local instrumented vehicle traffic density by counting the number of instrumented vehicles within their radio range, and vehicles also know their own speed. With the density and speed information, as an approximation vehicles can estimate the instrumented vehicle flow by applying the basic traffic stream model to instrumented vehicles only. As will be seen this flow estimate will be utilized in a vehicle's determination of the group membership of active forwarders. It should be recognized that only a rough approximation of vehicle traffic flow is provided by this procedure. Several factors contribute to potential estimation errors, including a failure to differentiate between vehicle traffic traveling in opposing directions, changing measurements on interrupted flow facilities (e.g. signalized intersections), and the reliance on the point observation of a single vehicle. However, for the determination of group membership this method has the distinct advantage of requiring no additional computation or data exchange. Improved traffic flow estimation would require the deployment of vehicle traffic monitoring infrastructure, such as the one managed by Traffic Management Center of Georgia Department of Transportation, or more frequent information exchange between vehicles. Both of these options may significantly increase the cost of the V2V network. Future research efforts are needed to explore the costs and benefits of this type of additional information.

3.4.2 Message Head Pair

Since the location of the message head changes over time, vehicles can only expect approximate message head location knowledge, and a vehicle's knowledge must be continuously updated. To accomplish this updating a small amount of information, the message head location and its generation time, collectively referred to as the *message head pair*, is placed in the message header. As the message is propagated among vehicles, it contains the message head pair. Every holder of a message maintains a message record containing the message head pair along with other information concerning this message.

The installed message head pair provides the best knowledge available to the message holder concerning the message head location. Based on the message head pair the message holder may state "as far as I know, the message head was in that location at that time".

Two messages differing only in the message head pair are two versions of the same message. One message version with message head pair $<l_i, t_i>$ is said to be newer than another message version with message head pair $<l_j, t_j>$ if: (1) l_i is closer to the destination region than l_j; or (2) $l_i = l_j$ but $t_i > t_j$. A vehicle always replaces its installed message head pair with the received newer information. Therefore obsolete/false information can be eliminated through data exchange.

A number of message holders, called message head candidates, are allowed to publish their own locations as the message head location, thus generating new message head pairs. Non-message head candidates can only learn from received messages. To reduce the publication and dissemination of inaccurate information about the message head location, efforts are required to constrain who can assume the role of message head candidate. A message holder assumes it is a message head candidate only when it determines it may be the real message head with some certainty. The remainder of this section defines the rules designed for a message holder to transition between a message head candidate and non-message head candidate. Suppose the current time is t_c, a vehicle's current location is l_c, and the message head pair is $<l, t>$, where l is the message head location and t is the generation time.

3.4.2.1 Non-message head candidate transition to message head candidate

The transition from non-message head candidate to message head candidate occurs when (1) a vehicle receiving a new message finds it is closer to the destination than what the message head pair in the message claims, or (2) a message holder passes the message location inside the message head pair and is traveling toward the destination region along the trajectory. As the time between the generation time in the message head pair and the occurrence of one of the two stated conditions increases, the likelihood of the message holder being the real message head decreases, as the message head pair is more likely to be obsolete. Thus a non-message head candidate is allowed to become a message head candidate only if one of the above two conditions occurs within t + T_1, where T_1 is a time period to be estimated using vehicle traffic information.

Specifying T_1 is a tradeoff between performance and message overhead. A large T_1 allows more message head candidates. But inaccurate information about the message head location may disrupt data dissemination and possibly result in higher message overhead. On the other hand, if no vehicle claims to be the message head candidate, vehicles will stop claiming to belong to the group of active forwarders (the reason for this to be described in the subsequent section), and the message has to rely on vehicle movement to reach the destination region. We estimate T_1 by limiting M_v, the average number of instrumented vehicles passing the message head location during the time period T_1. M_v is an indicator of how many vehicles are claiming to be message head candidates at that time. $M_v = T_1 * q = v_{thresh}$ where q is the vehicle traffic flow estimated, and v_{thresh} is a threshold value. T_1 can then be computed as $T_1 = \frac{v_{thresh}}{q}$. Simulation experiments for evaluating v_{thresh} are discussed later.

3.4.2.2 *Message head candidate transition to non-message head candidate*

For this transition there are two rules: (1) if the message head candidate leaves the trajectory or moves away from the destination region along the trajectory; or (2) if a message head candidate receives the same message with another message head pair $< l_n, t_n >$ where l_n is closer to the destination region than l_c, it becomes a non-message head candidate.

3.4.3 Data Exchange

In MDDV, every transmission is a local broadcast so that multiple receivers can receive the same message via one transmission. Data exchange is triggered by: the reception of new messages, messages with newer/older message versions, or the appearance of new neighbors. Transmissions triggered by new neighbors help overcome network partitioning. Transmissions triggered by the reception of new messages or newer message versions serve to quickly propagate messages or up-to-date message head location information. Transmissions triggered by the reception of messages with older message versions can eliminate false/obsolete message head information. If vehicles know the locations of neighbors, transmissions triggered by new neighbors or new messages only happen if there are neighbors closer to the destination region. This protocol is called the *full protocol*. The full protocol is designed to catch every forwarding opportunity. However the full protocol is also expensive in terms of message overhead so constraints are set on which part of the protocol a vehicle is allowed to execute at any instant. For this purpose, we design several transmission states for vehicles during which vehicles will execute different data exchange protocols.

A message holder can be in either one of two transmission states: the active state and passive state (these states are separate from the roles concerning message head candidacy described in the preceding section). A message holder in the active state runs the full protocol to actively propagate the message while a message holder in the passive state only transmits the message if it hears some older message version. Vehicles are in the active transmission state if they are close to the message head location. We limit those in the active transmission state both temporally and geographically. Given a message holder's installed message head pair <l, t>, its current location l_c and the current time t_c, it is in the active state if $t_c < t + T_2$ and l_c is within the distance L_2 from l, or else it is in the passive state as long as it remains in the forwarding trajectory and the lifetime of the message has not expired. Otherwise, the message will be dropped. T_2 and L_2 can be estimated using vehicle traffic information.

T_2 (temporally) and L_2 (geographically) define the group membership of those vehicles actively propagating the message during the forwarding phase. Specification of T_2 and L_2 is a tradeoff between data delivery performance and message overhead. We set $T_2 = T_1$ because message transmission with a delay larger than T_1 cannot yield new message head candidates (thus new message head pairs). We estimate L_2 by limiting N_v, the average number of instrumented vehicles in a road segment of length L behind the message head location. N_v is an indicator of message popularity. $N_v = L * \lambda = v_{thresh} \Rightarrow L = \frac{v_{thresh}}{\lambda}$ where λ is the vehicle density and v_{thresh} is the threshold value as before. L_2 is computed as $L_2 = (1 - \frac{t_c - t}{T_2}) * L$ where t_c is the current time and t is the generation time in the message head pair. In this way, L_2 becomes smaller as time elapses and equals 0 once the delay becomes T_2.

3.5 Optimistic Forwarding

Optimistic forwarding was first described by Chen and Kung [17]. But this work did not specify a detailed algorithm implementation. Here we present a detailed design of an *optimistic forwarding* algorithm.

Recall that *optimistic forwarding* approximates the ideal scenario by designating a message owner during the forwarding phase. Only the message owner is allowed to forward the message. The forwarding strategy is to transfer ownership to vehicles closer to the destination region whenever possible.

If vehicles are aware of the location of neighbors, *optimistic forwarding* can be designed as a variation of geographic forwarding. The current message owner holds the message until a new neighbor closer to the destination region along the trajectory appears. The message ownership is then transferred to that neighbor through a communication exchange.

If vehicles do not know the location of neighbors, the message owner would not know to whom to forward the message. To overcome this problem, we employ a contention based scheme that was introduced by Fübler [27] in designing geographical forwarding where the location of neighbors is either inaccurate or does not exist. The contention based scheme works as follows. The current message owner attaches its current location in the message and sends the message as a request for contention. Vehicles receiving the message compare their own locations with the location specified in the message. If their locations are closer to the destination region along the trajectory, they join the contention and set a backoff timer. If a contender does not hear the message again before its backoff timer expires, it assumes it wins the contention and becomes the new message owner. Otherwise it drops the contention. The backoff time is computed with $t_{backoff} = (1 - \frac{d}{r}) * t_{max}$ where d is the distance to the sender, r is the radio range, and t_{max} is the maximum backoff time. This means that a vehicle farther away from the sender sets a shorter backoff timer and is more likely to win the contention. Therefore the ownership is transferred in a way to make the greatest progress.

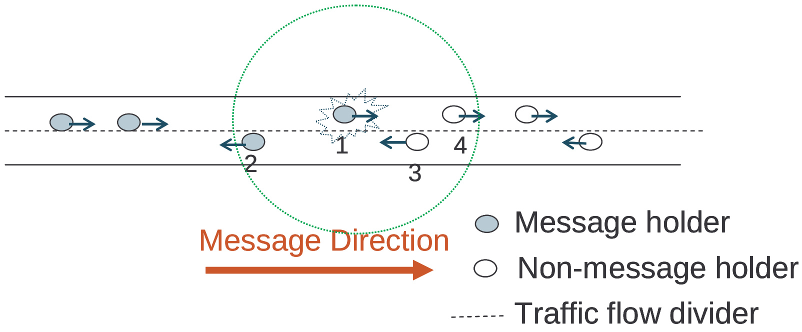


Figure 20: Contention-based scheme

One example is shown in Figure 20. Node 1 is the current message owner and sends out the message as a request for contention. Node 2, 3, and 4 receive the message almost simultaneously. Node 2 decides it is farther away from the destination than the sender (node 1) and will not participate in the contention. Node 3 and 4 determine they are potential candidates after comparing their locations with the location in the message and each sets a timer respectively. Node 4 has a shorter timer since it is farther away from node 1. After its timer expires node 4 wins the contention and becomes a new message owner. Node 4 will then send out the message again to try to forward the message even further. Upon receiving the message sent by Node 4, node 3 knows someone else closer to the destination than itself has assumed the message owner and thus will drop its current contention participation.

Details of the contention scheme for individual vehicles are represented in an event-driven processing algorithm shown in Algorithm 1. Part 1, 2 and 3 specify different conditions when a message owner sends out the request for contention. Part 3 states that a message owner transmits the message at least once every *opportunistic interval*, where the opportunistic interval is an algorithm parameter. This is designed to deal with the scenario in Figure 19(b) where a neighbor may become closer to the destination region along the trajectory due to movement. Part 4 specifies the operation when a message is received. Part 5 addresses the operation when the backoff timer expires.

1. A new message m is sent down from an application:
Set myself to be the owner of the message
Attach my location to the message
Send the message as the request for contention
Schedule a backoff timer of maximum backoff time
2. New neighbor appears:
For every stored message m for which I am the owner
Attach my location to the message
Send the message as the request for contention
Schedule a backoff timer of maximum backoff time
3. A stored message m is not transmitted/received for more than the opportunistic interval:
If(I am the owner)
Attach my location to the message
Send the message as the request for contention
Schedule a backoff timer of maximum backoff time
4. A message m is received:
If(I am on the trajectory and my location is closer to the destination region than the sender)
Cancel the previous backoff timer if any
Schedule a backoff timer based on the distance to the sender
5. The backoff timer for a message m expires:
If(I am the owner)
Continue to be the owner
else

Set myself to be the owner of the message
Attach my location to the message
Send the message as the request for contention
Schedule a backoff timer of max backoff time

Algorithm 1

3.6 Propagation Phase

During the propagation phase, the common behavior of MDDV and *optimistic forwarding* is relatively simple. A vehicle inside the destination region will transmit a message if it hears the message the first time or some new neighbors appear. The former serves to quickly propagate messages. The latter is to overcome network partitioning. The algorithm is represented with an event-driven processing (see Algorithm 2). In Algorithm 2 we use the backoff timer to reduce message transmission redundancy [61].

1. *A message m is received the first time:*
   If(I am inside the destination region)
   Schedule a random backoff timer
2. *New neighbor appears:*
   For every stored message m of which I am inside the destination region
   Transmit m
3. *The backoff timer for a message m expires:*
   Transmit m
4. *A message m is heard again before its backoff timer expires:*
   Cancel the backoff timer

Algorithm 2

3.7 Evaluation

In this section, we present an evaluation of the performance of the MDDV and *optimistic forwarding* algorithms. A simulation test bed described in Chapter 1.4 is used for the evaluation under realistic conditions. Even though metropolitan areas are usually covered with cellular networks, short-range V2V communications may be preferred for certain applications for the foreseeable future due to cellular cost and network capacity issues.

3.7.1 Simulation Methodology

We use a vehicle traffic model simulating the morning rush hour traffic for the northwest quadrant of Atlanta, Georgia (see Figure 2). We implement opportunistic forwarding algorithms (MDDV and *optimistic forwarding*) in a data service layer between the application and the transport layer. To study the best performance achievable using store-carry-forward schemes, we also implement an **ideal** algorithm, where perfect knowledge is assumed (i.e. the identity of the message head is known to all vehicles) and only the message head may transmit the message during the forwarding phase. All three algorithms exploit the local broadcast nature of wireless transmission to reduce message overhead and utilize the propagation phase algorithm described in Chapter 3.6. In this evaluation, for all three algorithms (MDDV, *optimistic forwarding*, and ideal algorithm) it is assumed vehicles do not know the exact location of neighbors. Dissemination performance of each algorithm is also evaluated under various penetration ratios. In these evaluations, vehicle instrumentation reliability is not taken into account, although it is clearly recognized as a critical aspect to be incorporated into future research.

Each simulation run is specified by a vehicle mobility trace, a penetration ratio, and a message propagation algorithm. For each simulation run, a fixed number of messages are disseminated. The same set of messages is disseminated by each of all three algorithms respectively. Each message is sent from a randomly chosen origin (i.e. point source on a surface street) to a randomly chosen destination (i.e. a surface street segment). The message source is always constrained to be outside of the destination region. The average road distance from the source to the destination region is approximately 6.5km. Since each message may have to traverse road segments with different vehicle traffic conditions (including traffic lights and signs), a very challenging case for data dissemination is created. The message expiration time is set as 480 seconds. The simulation time is 650 seconds, of which the first 150 seconds are used for initialization, and messages are generated between 150s and 170s. We use a two ray wireless signal propagation model since terrain data are not available for these experiments. Every instrumented vehicle is equipped with an omni-directional antenna. The radio range is set at 250m, limiting data hops primarily along the roadway network. The MAC protocol is 2Mbps 802.11 DCF broadcast. It should be noted that the proposed approaches do not require any specific short range wireless technologies. We choose 802.11 because DSRC is based on 802.11a. After some trials, we set the opportunistic interval in *optimistic forwarding* as 4 seconds because this value showed a good cost/performance tradeoff. Results presented in this section are the aggregation/average of 4 runs with different random seeds.

3.7.2 Baseline Scenario

In assessing the performance of the MDDV and *optimistic forwarding* algorithm it is instructive to establish a baseline that reflects optimal message forwarding under the given traffic conditions. We demonstrate this by studying the ideal algorithm, described in the preceding section.

Algorithm performance is measured in terms of **dissemination delay**, defined as the time for a message to reach the destination region. As stated previously, a message will arrive at the destination region faster if more data forwarding opportunities are captured; otherwise the message has to rely increasingly on vehicle movement to reach the destination region, a significantly slower method of message forwarding. If vehicle movement is the principal mean of message dessemination, in the best case, the message originator travels to the destination region. Suppose the message originator moves at a speed of 50km/h (about the average speed of vehicles on surface streets in our model) for a distance of 6.5km (the average road distance of moving trajectories), it will take about 480 seconds for the message to reach its destination region. Furthermore, vehicles have to stop on traffic lights and signs. And because there is often no vehicle moving directly from the message origin location to the destination region, at times messages have to wait to be transferred from one vehicle to another. All these factors will contribute to a much longer dissemination delay if vehicle movement is used as the principal way of message dissemination. Next we plot the cumulative distribution of dissemination delays of all messages when the ideal algorithm is used.

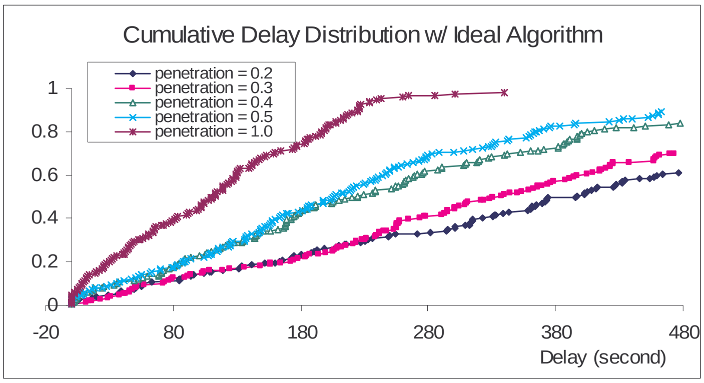
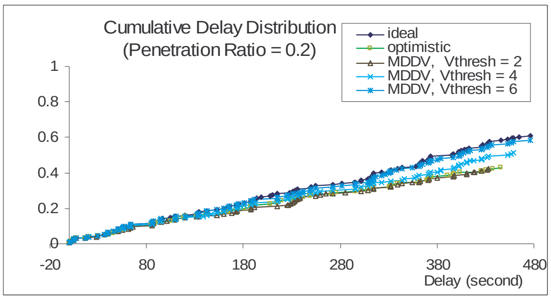


Figure 21: Cumulative distribution of dissemination delays of all messages for various penetration ratios w/ the ideal algorithm

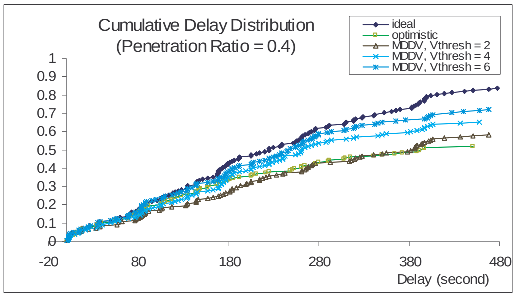
As shown in Figure 21, message forwarding performance improves as the penetration ratio increases, since more forwarding opportunities exist. In interpreting Figure 21 and several figures that follow, for a given delay, higher cumulative distribution values imply better performance. When the penetration ratio is 1.0, almost all messages reach their destination regions within their 480-second lifetime with only a few exceptions with an average delay of about 110 seconds, significantly faster than relying principally on vehicle movement. Dissemination delays are somewhat uniformly distributed, as demonstrated by the nearly linear relationship on the cumulative delay distribution. This is likely a reflection of the random selection of source locations and destination regions, resulting in forwarding trajectories of different lengths and forwarding paths traversing road segments with different vehicle traffic conditions. However, most messages require at least 10 seconds reaching their destination regions. Even at the maximum possible penetration ratio, 1.0, 90% of the dissemination delays are more than 10 seconds. Where an end-to-end connection exists, the delay should be much less than 10 seconds, as the actual single V2V message hop time is in the order of milliseconds. Therefore, in the scenarios studied, any algorithm assuming end-to-end connection (e.g., DSR) will fail to deliver most messages. Instead a store-carry-forward scheme is necessary for reliable message delivery as such schemes work regardless of whether end-to-end connections exist.

3.7.3 Performance Comparison

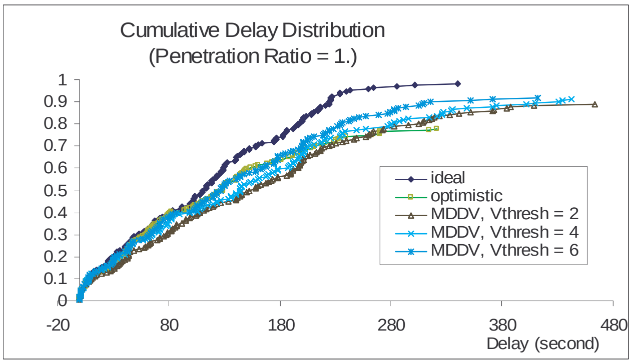
We compare the three algorithms under different penetration ratios. Since they all utilize the same algorithm during the propagation phase, we focus on comparing their performance during the forwarding phase.



(a) penetration ratio = 0.2



(b) penetration ratio = 0.4



(c) penetration ratio = 1.0

Figure 22: Performance comparison

Figure 22 contains plots of the cumulative delay distribution of all messages for the ideal algorithm, *optimistic forwarding*, and MDDV with different v_{thresh} for three different penetration ratios (0.2, 0.4, and 1.0). Remember v_{thresh} is the threshold value in MDDV used to define the group of vehicles actively forwarding the message. For MDDV, performance with v_{thresh} =2, 4, and 6 is plotted. Through these plots it is readily verified that the performance of MDDV and *optimistic forwarding* improves as the penetration ratio increases, a result of the increased forwarding opportunities.

As expected the ideal algorithm outperforms the other two algorithms. Where vehicle densities are high or the forwarding trajectory is short all three algorithms are capable of providing similar low delay values (as demonstrated by the nearly merging of cumulative delay distribution when the delay is low). However, the cumulative delay distribution increases at a greater rate for the ideal algorithm than both MDDV and *optimistic forwarding*, reflecting the ability of ideal algorithm to better take advantage of message forwarding opportunities under lower vehicle density and longer trajectories conditions. MDDV and *optimistic forwarding* rely more on vehicle movement for message dissemination because the message head in MDDV or message owner in *optimistic forwarding* may be lost, resulting missed forwarding opportunities.

For the MDDV algorithm, the performance improves and approaches the ideal algorithm as v_{thresh} increases, a direct result of the increasing active forwarder group size. The disparity between the ideal algorithm and MDDV performance varies because our method of estimating vehicle traffic conditions is inaccurate. As stated in Chapter 3.4.1, the estimated vehicle traffic flow does not account for the uneven distribution of vehicles along the road. As a result, at times the estimated vehicle traffic flow is larger than the actual, resulting in the number of message head candidates or the group of active forwarders smaller than necessary. Under some circumstances the message head is lost if no vehicle claims to be the message head candidate, the message will stop being forwarded, and the message will have to rely on vehicle movement to reach the destination region. This will result in higher delays and reduced performance. We have also tried to fix MDDV parameters T_1, T_2, and L_2 but found this approach does not work well due to the uneven distribution of vehicle traffic in this model. We expect further MDDV improvements could be achieved with more accurate vehicle traffic estimation.

Optimistic forwarding performs much worse than the ideal algorithm as a message owner may drive out of the forwarding trajectory, after which that message must rely on vehicle movement to reach its destination region. It comes with no surprise that with large enough v_{thresh}, MDDV can always outperform *optimistic forwarding* significantly due to the message replication. In addition, if vehicle instrumentation reliability is taken into account, we expect *optimistic forwarding* will perform much worse due to single point of failure.

3.7.4 Message Overhead

Message overhead is defined as the number of message transmissions during the forwarding phase. Figure 23 provides a rough comparison of the message overhead of these 3 algorithms. The message overhead of MDDV and *optimistic forwarding* is shown as normalized against that of the ideal algorithm.

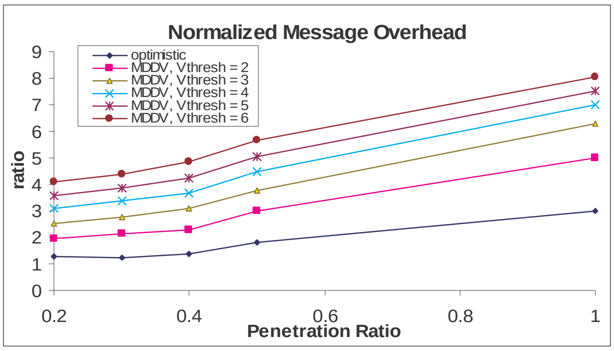


Figure 23: Normalized message overhead

Let us first compare *optimistic forwarding* with the ideal algorithm. With *optimistic forwarding*, some messages stop being forwarded during the forwarding phase because message owners drive out of their forwarding trajectories, resulting in reduced message overhead, but also reduced performance. With the ideal algorithm, there is always a message owner actively forwarding the message, potentially leading to more message transmissions, but better performance. However, a significant volume of *optimistic forwarding* message transmissions result from the requirement that a message owner transmits the message at least once every opportunistic interval. We have attempted to balance the size of the opportunistic interval and overall performance. As a result, *optimistic forwarding* is shown to incur a little more overhead, although its performance is much poorer than the ideal algorithm as shown in the preceding section.

As expected, MDDV incurs several times higher overhead than the ideal algorithm and *optimistic forwarding* due to message replication. Also, as the v_{thresh} increases, the number of vehicles actively forwarding increases, leading to higher overhead.

3.8 Discussion

3.8.1 Design Space

V2V communications certainly have limitations, as shown through the evaluation of the ideal algorithm in the Chapter 3.7.2. Interactive applications (e.g., telephony, gaming) require fast data delivery and thus cannot tolerate delays on the order of tens of or even hundreds of seconds. To be successful over a network these types of applications will require extra network support (e.g., roadside relays) designed to supplement V2V communications [88]. However, multi-hop V2V communications are suitable for information services that can tolerate some data delay and loss (e.g., traffic and tourist information), implying there are many useful services that may be satisfied by inexpensive V2V communications. We also show that practical algorithms (e.g., MDDV) can be designed to approximate the message delay performance of the ideal algorithm, although these algorithms incur higher overhead.

The two enhanced opportunistic forwarding algorithms proposed here, MDDV and *optimistic forwarding*, have their own merits and shortcomings. MDDV has been shown to outperform *optimistic forwarding*. However, *optimistic forwarding* incurs much less overhead than MDDV and thus a more attractive option where vehicles are less likely to move away from the specified data forwarding trajectory (e.g. few alternate roads), and vehicle instrumentation reliability is not a significant concern. Where these conditions are not satisfied MDDV is the preferred alternative.

In our current design the data source specifies the global data dissemination information, e.g., the destination region and forwarding trajectory. Intermediate vehicles have the flexibility of tuning local algorithm parameters based on their own knowledge. This arrangement has the benefit of easing coordination between vehicles, somewhat alleviating the difficulty typically experienced in coordination between peers in a distributed system with high dynamics. Building upon this flexibility, one area of future research is to allow intermediate vehicles more flexibility in manipulating messages. For example, intermediate vehicles may specify a better forwarding trajectory, change the destination region, or aggregate multiple messages based on application semantics. However, this increased flexibility will likely create new peer coordination issues. For example, if one peer changes the forwarding trajectory other peers may not be aware of, or agrees with, such a change. This may result in data propagation along multiple forwarding trajectories simultaneously.

Additional future efforts may also consider the possible expansion from only one forwarding trajectory to multiple diverse forwarding trajectories, in an attempt to improve performance and reliability. Additional trajectory specification may incorporate criteria other than shortest path, e.g., instrumented vehicle density and expected network partitioning.

3.8.2 Extension

The data dissemination services mentioned in Chapter 3.2 can be easily implemented with extensions to the design for geographical-temporal multicast, as discussed next.

Unicast. Unicast with precise location can be handled by the above design in the forwarding phase. Once the message reaches the destination location, the message will not be forwarded any longer. The algorithm for unicast with approximate location is an extension to the algorithm for unicast with precise location, beyond which intermediate vehicles must estimate the location of the destination as time elapses and determine whether the message has reached the destination.

Scan. It can be easily seen that the above design scans over the forwarding trajectory when delivering the message to the destination region. Scan can be implemented as a special case of unicast with precise location. The destination location is set to the other end of the region to be scanned. The moving trajectory is set as going through the region.

Anycast. The algorithm for anycast is similar to the algorithm for scan. However the message does not have to traverse the entire destination region. A reply from any intended receiver will quench further propagation.

Multicast. The scenario that the source is inside the destination region is a special case of geographical-temporal multicast. When there are multiple disjoint destination regions, the moving trajectory can be specified as a dissemination tree with the location of the message originator as the root. Some implementation issues need to be addressed concerning branches.

3.9 Summary

In this chapter, we have shown that opportunistic forwarding is a viable approach for data dissemination in a partitioned and highly mobile vehicular network for applications that can tolerate some delays and data loss. Messages are forwarded along a predefined trajectory geographically closer to the destination. Since end-to-end connectivity cannot be assumed in a V2V network, intermediate vehicles must buffer, carry and forward messages opportunistically. We present a generic methodology to design enhanced opportunistic forwarding algorithms. We then describe two algorithms, MDDV and *optimistic forwarding*, as derived from this generic methodology. Our evaluation shows the limitation of multi-hop V2V communications for data dissemination due to uncertainties inherent in V2V networks (e.g., high vehicle mobility, insufficient vehicle density, network partitioning, etc). Multi-hop V2V communications are best suitable for information services that can tolerate some data delay and loss. Practical algorithms can be designed to approximate an ideal algorithm that assumes perfect knowledge, although these algorithms incur higher overhead.

CHAPTER

4 INFRASTRUCTURE-BASED VEHICULAR NETWORKS

4.1 Introduction

Chapter 1.3 gave an overview of infrastructure-based vehicular networks. Previous research in vehicular networks have focused on infrastructure-less V2V communications, e.g., see [79, 87, 89]. This deployment offers the benefit of low cost and easy deployment, and is appropriate for some localized applications (e.g., collision avoidance), but fails to provide reliable communication services, especially where the density of instrumented vehicles is low. Also a pure V2V network is a standalone network and cannot provide access to external online resources such as the Internet. The motivation to build infrastructure-based vehicular networks is to provide reliable broadband communication services, access online resources, communicate with other people, and access local services (e.g., traffic information) not residing within vehicles. A number of existing wireless technologies can be exploited to creating vehicular networks, including WWAN, WLAN using roadside access points, and ad hoc networks using short-range communication between vehicles. These technologies offer different tradeoffs in cost and performance.

Today there is a major shift from core to edge networks (an edge network is a network located on the periphery of a centralized network. The edge network feeds the central, or core, network.) As a result we have seen a rapid growth and evolution of

WWAN and WLAN in providing network connectivity and Internet access for mobile users. While WWAN is aiming to provide ubiquitous coverage (e.g., Verizon Wireless Broadband Access), WLAN offers high bandwidth in a limited area (e.g., T-Mobile HotSpots). City-operated WLAN broadband networks are now under discussion in Chicago, Philadelphia, Las Vegas, New York, and San Francisco. There is also extensive research on combining the two technologies to leverage the high capacity of the WLAN and wide coverage of the WWAN [33-35, 49, 52, 66, 92]. However, comparatively little work has been done on exploring wireless infrastructure for travelers inside vehicles. Research is necessary to: (1) evaluate communications architectures to identify those that are best suited for providing high bandwidth communications to travelers, (2) examine design options and tradeoffs that arise in realizing such networks, and 3) quantitatively assess alternate approaches and evaluate their performance and reliability under realistic traffic and environmental conditions. This chapter presents an initial investigation to address the above issues and seeks to motivate more work in this area.

Following the discussion in Chapter 1.3, urban and rural areas may deploy different vehicular networks based on cost-benefit tradeoffs. In urban areas, just as other wireless services, it is reasonable to assume that wireless infrastructures are deployed to provide ubiquitous connectivity in a cost-effective way. V2V communications can also be used for direct inter-vehicle information exchange. In rural areas, it might be more economical to rely on V2V communications supplemented by limited infrastructures in hot spots. In this study, we systematically identify and evaluate several feasible wireless network architectures to provide ubiquitous connectivity in urban areas. But some of our analysis also applies to hot spots in the “rural network”.

The infrastructure may leverage various wireless technologies, e.g., WWAN and WLAN, to work together in a seamless fashion. In the “urban network” of previously presented in Figure 1, WWAN base stations and WLAN access points are all connected to a backbone through wired links or fixed broadband wireless links (e.g., WiMAX [3]) which itself is connected to the Internet. Users can access the WWAN directly anywhere and anytime. However, due to its limited bandwidth and high cost, WWAN by itself may not be able to meet demands that are placed upon it. For example, if a 2Mbps WWAN covers an area containing 100 instrumented vehicles, a single vehicle can only obtain an average data rate of less than 2Kbps. This is clearly not enough for delivering rich content media. WLAN access points can be placed along the road to provide high-bandwidth and low-cost communication services. A vehicle can access a roadside WLAN access point either directly or through the relay of other vehicles. A vehicle may be equipped with multiple wireless interfaces (e.g., cellular, 802.11x, DSRC etc.) attached to devices that are interconnected through an internal network (e.g., Ethernet or Bluetooth). These devices cooperate to provide travelers the required communications services. In this way one can create a “mobile intranet” on the road.

Depending on the wireless network building blocks (WWAN, WLAN last-hop and multi-hop WLAN) in an infrastructure, there are several different design options in providing ubiquitous connectivity. Hsieh et al. [33] proposed the following classification:

WWAN last-hop. Cellular-based WWANs are deployed to cover large areas. Since cellular networks have covered most metro areas, this design has already been realized in these areas. However, as discussed before, conventional cellular networks are unable to support a large number of high-bandwidth users. Deploying high bandwidth cellular networks would be too expensive or infeasible with existing technologies. In addition, some “dead” spots may arise that are difficult to cover.

WLAN last-hop. High-speed WLAN access points are placed along the road. Users can experience high throughput within the WLAN coverage area with much lower cost than the WWAN service. Due to the limited reach of WLAN, covering major roads in a metro area requires a large number of WLAN access points. Even though the initial installation cost may not be too high (see [2] for a cost estimation), significant additional costs may well accrue with ongoing maintenance and upgrades for a large number of WLAN access points [4]. Again, some “dead” spots may arise.

Multi-hop WLAN. This is an extension of the previous design. Ad-hoc V2V communications between vehicles are exploited to extend the reach of WLAN access points. Thus, the number of WLAN access points used to provide full coverage can be largely reduced. In addition, vehicle forwarding can be used to reduce “dead” spots. But connectivity cannot be guaranteed due to the uncertainty of ad hoc communications.

WWAN last-hop + WLAN last-hop. The WWAN covers the entire area while WLAN access points are placed in some spots. In the overlapped areas of the WWAN and WLAN, vehicles can choose to access the WLAN or WWAN or both. Outside the WLAN coverage, vehicles can only access the WWAN. Compared with the WWAN last-hop approach, this design offers higher capacity and can support more users. Compared with the WLAN last-hop approach, this design can reduce “dead” spots and provide more reliable services. This design is especially appealing in places where the cellular network already exists.

WWAN last-hop + Multi-hop WLAN. This can be viewed as an extension of the previous design. Vehicles can access the WLAN either directly or through the forwarding of other vehicles. The maximum number of allowed forwarding hops is usually determined by delay requirements, data traffic load, vehicle connectivity, and other factors (e.g., routing overhead), as we will address later. Multi-hop forwarding extends the coverage of the WLAN and reduces the number of required WLAN access points. The vulnerability of multi-hop paths is mitigated because the WWAN can provide the connectivity when the path is absent.

4.2 Evaluation

In this section, we present quantitative evaluation to verify the above analysis and also examine some other issues of interest. When evaluating vehicular networks we must take vehicle traffic conditions into account. Following the discussion in Chapter 1.4, we use both statistical analysis and simulation for large-scale network infrastructure evaluation. In this section, we first introduce our evaluation models. We then assess individual building blocks (WWAN, WLAN last-hop and multi-hop WLAN), and show their relevance and interaction when integrated into a network infrastructure.

4.2.1 Test Scenario

We introduce the evaluation with reference to a specific test scenario. In particular, we study constructing a wireless network infrastructure providing **continuous** communication services for vehicles on the I-75 freeway in a segment of approximately 11 km in the northwest quadrant of Atlanta, GA. We choose freeways because they normally have a high concentration of vehicles (thus posing high demand for communication services), and are most likely to be cost-effective. This is the same reason commercial cellular networks cover major freeways. We expect the traffic conditions on I-75 to be representative of that in many metro areas in the U.S.

Each instrumented vehicle is equipped with two network interfaces (if necessary): one for the WWAN and another for the WLAN. The WLAN and WWAN are operating on different frequencies and are assumed not to interfere with each other. The WLAN radio range is set at 250m for both V2V and V2R communications based on our communication performance measurement study to be introduced in the next chapter. Our study focuses on upstream data traffic[3]. Every instrumented vehicle on the freeway generates a constant UDP flow[4] to the infrastructure. We consider a packet as having reached its destination once it reaches any WWAN base station or WLAN access point. We explore different instrumented vehicle densities by varying the penetration ratio (fraction of instrumented vehicles in the vehicle traffic stream).

To perform evaluations, we must first define how vehicles access network resources. In this study, we assume data traffic needs to be transmitted immediately, even through some applications (e.g., email) can tolerate some intentional delay. Because WWAN is considered scarce and expensive, we consider a set of access policies where vehicles always try to access the WLAN first. We assume vehicles employ a scheduling policy to decide how to transmit each packet. Based on the mechanism that a vehicle uses to decide whether the WLAN is accessible, one can envision two policies:

1. **Deterministic policy**. A vehicle sends a packet through the WLAN if the vehicle has a path of fewer than m hops to a WLAN access point. To implement this policy

[3] We realize that user data traffic is typically asymmetric. There are some transportation applications principally using upstream communication, e.g., reporting vehicle traffic information to a traffic management center. Certain

each vehicle must know whether such a path exists. One approach is to exploit WWAN to inform vehicles of the network topology, as suggested by [33, 52]. With this approach vehicles periodically send updates of their locations to a central coordinator through a WWAN link and the aggregated information will be broadcasted back to the vehicles. Another approach is to have vehicles run a localized routing protocol. Due to vehicles' high mobility, both approaches will likely incur non-negligible overhead.

2. **Statistical policy**. We assume vehicles know the location of WLAN access points. Statistics about vehicle traffic conditions can broadcast from the infrastructure to vehicles through a shared WWAN link. Vehicles use this information to compute P_m, the probability that they are connected to the closest WLAN access point in fewer than m hops using an analytical model to be introduced momentarily, based on which vehicles can determine which network to use. For example, it might use the WLAN if the probability is above some threshold p_t. In this way vehicles only need to maintain one-hop neighborhood information. Intermediate vehicles will forward the packet to the next hop if one exists (e.g., based on position), or otherwise transmit it through the WWAN so that data will not be lost. Since this is a statistical scheme, there will be some *missed* opportunities (i.e., there is a path to a WLAN access point, but the vehicle determines the WLAN is not accessible) and *false* opportunities (i.e., there is no path to the WLAN, but data is sent to the next hop). We have observed both missed and false opportunities in our simulation

commonly used applications (e.g., vehicle tracking, video/picture uploading) also generate significant upstream data traffic.

[4] We do not use adaptive TCP flows because we want to study system behavior under different data traffic loads.

study. If p_t is high, false opportunities are infrequent but miss opportunities are probable. If p_t is low, missed opportunities are infrequent but false opportunities are probable. Properly setting p_t to balance these two factors is an area of future research.

Our evaluation focuses on the deterministic policy. However, we will discuss how the statistical policy may affect these results, when applicable.

4.2.2 Analytical Models

Our analytical models utilize certain vehicle traffic information that can be readily obtained. As described in Chapter 1.5.3, the three most important characteristics of vehicle traffic are flow *q* (vehicles/h), speed *u* (km/h) and density λ (vehicles/km). The average values of these quantities can be approximately related with the basic traffic stream model $u = q / \lambda$. Transportation authorities regularly publish traffic information[5], e.g., traffic flow and traffic speed, with which we can approximately derive the vehicle density using this basic traffic stream model.

We wish to compute the probability at an instant in time that a vehicle can communicate with a roadside WLAN access point using at most m hops through other instrumented vehicles. That is, we wish to compute $P_m(x)$: the probability that a vehicle with a distance x from a roadside WLAN access point can reach the access point in at most m hops. Intuitively $P_i(x) \geq P_j(x)$ when $i > j$. The derivation requires certain assumptions concerning vehicle traffic conditions.

We model the road as one-dimensional since the width of the road is relatively small compared to the radio range. To simplify the problem, it is assumed that spatial

[5] Examples are Georgia Traffic Management Center (TMC) operated by Georgia Department of Transportation.

positions of vehicles can be modeled as a one-dimensional Poisson point process with the instrumented vehicle traffic density λ. A Poisson process captures the randomness of vehicle locations. Usually vehicle locations are not totally random, e.g., a vehicle must keep some distance from the vehicle in the same lane that is in front of it (that is where the simulation model comes in). When considering all the free-flow vehicles in a multi-lane road such as a freeway in a metropolitan area, vehicle locations can be approximated as Poisson process [70].

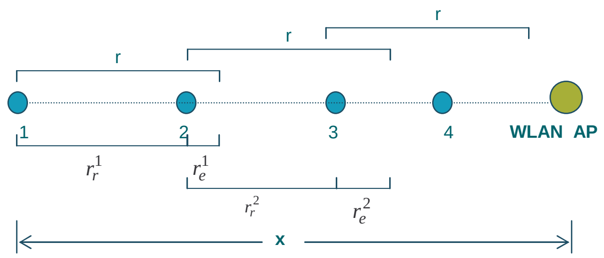


Figure 24: One-dimensional forwarding model

The situation is illustrated in Figure 24. We compute $P_m(x)$ for node 1, which has a distance x to the WLAN access point. We first review the analysis by Cheng and Robertazzi [18] that forms the basis for our model. In the first hop, the rightmost node that node 1 can reach using a transmission is node 2. Without loss of generality, only the receiving station (node 2) most distant from the transmitter (node 1) will be considered. Node 2 is at a distance r_r^1 (advancement) from node 1. r_e^1 is the empty gap of the current hop. $r = r_r^1 + r_e^1$ where r is the radio range.

Conditioned on the event that there is at least one vehicle within the radio range of node 1 toward the access point, given the Poisson process assumption, $f(r_r^1)$, the PDF of r_r^1, is

$$f(r_r^1) = \frac{\lambda e^{-\lambda(r-r_r^1)}}{1-e^{-\lambda r}} \quad 0 < r_r^1 < r \tag{4-1}$$

where λ is the density of instrumented vehicles.

$f(r_e^1)$, The PDF of r_e^1, is

$$f(r_e^1) = \frac{\lambda e^{-\lambda r_e^1}}{1-e^{-\lambda r}} \quad 0 < r_e^1 < r \tag{4-2}$$

The next hop can be considered as originating from node 2. In general, suppose r_r^i is the advancement toward the access point in the *i*th hop and r_e^i is the empty gap in the *i*th hop, we have $r = r_r^i + r_e^i$.

Conditioned on the event that there is at least one receiver toward the access point in the *i*th hop, the PDF of r_r^i is

$$f(r_r^i) = \frac{\lambda e^{-\lambda(r-r_r^i)}}{1-e^{-\lambda(r-r_e^{i-1})}} \quad r_e^{i\text{-}1} < r_r^i < r \tag{4-3}$$

The PDF of r_e^i is

$$f(r_e^i) = \frac{\lambda e^{-\lambda r_e^i}}{1-e^{-\lambda(r-r_e^{i-1})}} \quad 0 < r_e^1 < r - r_e^{i-1} \tag{4-4}$$

Next we compute $P_m(x)$.

Define $fr(\xi) = \lambda e^{-\lambda(r-\xi)}$

For m = 1,

$$P_1(x) = \begin{cases} 1 & 0 < 0 \le r \\ 0 & \text{else} \end{cases} \tag{4-5}$$

For m = 2, $P_2(x)$ is the probability that the first i (i<2) hop reaches a node that is located at a distance less than r from the access point. We have

$$P_2(x) = \begin{cases} 1 & 0 < x \le r \\ (1-e^{-\lambda r})\int_{x-r}^{r} f(r_r^1)dr_1 = \int_{x-r}^{r} fr(r_1)dr_1 & r < x \le 2r \\ 0 & \text{else} \end{cases} \qquad (4\text{-}6)$$

where $(1-e^{-\lambda r})$ is the probability that the first hop exists.

For m = 3, $P_3(x)$ is the probability that the first i (i<3) hops reach a node that is located at a distance less than r from the access point. After some manipulations, we have

$$P_3(x) = \begin{cases} 1 & 0 < x \le r \\ \int_0^r fr(r_1) * \left\{ \begin{matrix} 1 & x - r_1 - r \le 0 \\ \int_{\max(r-r_1,\, x-r_1-r)}^{r} fr(r_2)dr_2 & \text{else} \end{matrix} \right\} dr_1 & r < x \le 3r \\ 0 & \text{else} \end{cases} \qquad (4\text{-}7)$$

In general, $P_m(x)$ is the probability that the first i (i<m) hops reach a node that is located at a distance less than r from the access point. After some manipulations, we get

$$P_m(x) = \begin{cases} 1 & 0 < x \le r \\ \int_0^r fr(r_1) \cdots \left\{ \begin{matrix} 1 & x - r - \sum_{n=1}^{m-3} r_n \le 0 \\ \int_{r-r_{m-3}}^{r} (fr(r_{m-2}) \left\{ \begin{matrix} 1 & x - r - \sum_{i=1}^{m-2} r_i \le 0 \\ \int_{\max(r-r_{m-2},\, x-r-\sum_{i=1}^{m-2} r_i)}^{r} fr(r_{m-1})dr_{m-1} & \text{else} \end{matrix} \right\}) dr_{m-2} & \text{else} \end{matrix} \right\} \cdots dr_1 & r < x \le m \cdot r \\ 0 & \text{else} \end{cases}$$

(4-8)

Based on this we compute $P_{d,m}$, the probability that a vehicle is connected to the WLAN in at most m hops when driving between two WLAN access points separated by a distance of d. This is equal to the probability that the vehicle is connected to either one of the two access points. Suppose the location of the vehicle is uniformly distributed among theses two access points, it can be seen that

$$P_{d,m} = \int_0^d (1-((1-P_m(x))(1-P_m(d-x)))\frac{1}{d}dx = \int_0^d (P_m(x)+P_m(d-x)-P_m(x)P_m(d-x))\frac{1}{d}dx \quad (4\text{-}9)$$

$P_{d,m}$ is also the average percentage of vehicles between these two WLAN access points connected to the WLAN in at most m hops at any instant, and the percentage of time a vehicle is connected to the WLAN while driving between these two WLAN access points.

To obtain numerical results, we must specify the numerical values of model parameters. The wireless communication range is chosen as 250m, a typical range of 802.11b. To be able to compare the analytical results with the results predicted by simulation models when necessary, the vehicle traffic parameters (e.g., speed and density), as input to analytical models, are derived from simulation traces. The length of the multi-lane freeway studied is 11km and the average number of vehicles within this section of the road is 1800 (Including both instrumented and non-instrumented vehicles). However, one can easily use these models to explore conjectures and projections by examining other parameter values.

4.2.3 Simulation Settings

The simulation study uses a test bed combining two simulation packages, CORSIM and QualNet, as discussed in Chapter 1.4. The CORSIM model used for this study simulates the vehicle traffic in the morning rush hours in the northwest quadrant of Atlanta.

To be able to study the network infrastructure, we must specify its configuration. The WLAN operates in DCF mode of IEEE802.11b at 11Mbps channel data rate. All the

WLAN interfaces operate in the same channel. Some proposals [12] suggest exploiting multiple orthogonal channels in multi-hop networks. However, complicated channel synchronization is required between neighboring nodes for these schemes. Each WWAN cell provides an upstream link, which is shared by all users in the cell using time multiplexing. Instead of modeling a specific WWAN standard, we employ a generic WWAN model[6]: the base station applies a round robin scheduling policy to deliver the head-of-queue packet from each vehicle, e.g., one packet from vehicle 1, one packet from vehicle 2, etc. Only the deterministic policy mentioned above is simulated. Multi-hop WLAN routing is not modeled but computed centrally.

The generation time of each data packet is slightly randomized to avoid lock step behaviors. The data packet size is set at 1KB. Once a packet is generated, if applicable, the vehicle decides whether to send it to the WLAN or WWAN. The WLAN will be used if there is a path of less than the **hop limit** to a WLAN access point and then the packet will be unicast to the next hop. Otherwise, the packet will be stored in a local WWAN pending queue and wait to be delivered to a WWAN base station. Each vehicle maintains a WWAN pending queue of 8 packets. The head of the queue will be dropped once the queue is full. By varying the hop limit allowed to access WLAN access points, we can model different design options. For example, if the hop limit is 1, we are modeling the WLAN last-hop.

The simulation time of each run is 100 seconds. The simulation results presented in this section is the aggregation of 4 runs using different random number seeds.

[6] We acknowledge that our WWAN model is a simplification. However we believe this does not undermine our results since it is not our intention to compare WWAN technologies in this study.

4.2.4 *Required WWAN Capacity*

Chapter 4.1 claimed that WWAN capacity is often insufficient to support the capacity demand of the instrumented vehicles under coverage. In this section, we provide some numerical results of the required WWAN capacity for different network architectures.

For a WWAN last-hop architecture, the WWAN data traffic load is the sum of data traffic of all the instrumented vehicles in the covered area. Suppose the average number of instrumented vehicles in the studied area is n, the per vehicle data rate is b, the average WWAN data traffic load l_{WWAN} is

$$l_{WWAN} = n \cdot b \tag{4-10}$$

For a WWAN+WLAN hybrid architecture, the WWAN data traffic load is the sum minus the data traffic assigned to the WLAN since the WLAN is preferred. If adjacent WLAN access points are separated by a distance of d, the average WWAN data traffic load is

$$l_{WWAN} = (1 - P_{d,m}) \cdot n \cdot b \tag{4-11}$$

when the deterministic policy is employed. If the statistical policy is applied, the WWAN data traffic load should be higher since all of the missed and false opportunities eventually become WWAN traffic.

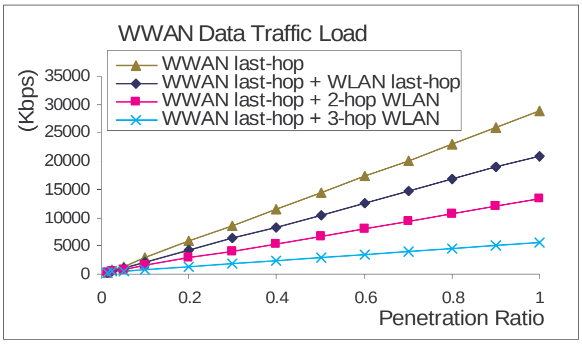


Figure 25: Average WWAN data traffic load for WWAN last-hop and hybrid WWAN+WLAN architectures with hop limit equals 1, 2, and 3

Figure 25 shows some numerical results for the studied area where the data traffic rate per vehicle is assumed as 16Kbps, a quite modest value. Seven WLAN access points with equal distance between neighbors are used in hybrid architectures. One access point is placed on each end of the studied road segment. The deterministic policy is employed. Without WLAN, the aggregated WWAN data traffic load can reach as high as 28.8 Mbps. This is well above the capacity of most existing commercial cellular networks, readily verifying that the WWAN last-hop architecture is normally not a feasible design option to meet the communication demands. A hybrid architecture of WWAN and WLAN increases the system capacity and reduces the WWAN data traffic load as some data traffic is diverted to WLAN, especially with multi-hop forwarding. When the hop limit is 3, the average WWAN data traffic load with 100% penetration ratio is about 5.6 Mbps. Three WWAN cells, each with a 2 Mbps link and covering about 1/3 of the studied area, can provide the required capacity.

4.2.5 WLAN Coverage and Average Data Traffic Load

Users may access a WLAN access point either directly or through the forwarding of other vehicles. The preceding section has shown that multi-hop forwarding also alleviate the WWAN traffic load in a hybrid architecture. Intuitively multi-hop forwarding can extend the coverage of an access point and thus increase its data traffic load, but only if the vehicle density is sufficiently high. In this section, we will examine whether instrumented vehicles are sufficiently dense in the area studied. We first quantify the coverage and data traffic load of an access point under different hop limits.

Let C_m be the maximum length of the road segment on one side of an access point within which vehicles can access the access point using at most m hops. C_m is a random variable. We quantify the coverage of a WLAN access point with c_m, the sum of expected values of C_m on both sides of the access point. If the deterministic policy above is used,

$$c_m = 2 \cdot E(C_m) = 2 \cdot \int_0^{\infty} P_m(x)dx \qquad (4\text{-}12)$$

where $E(C_m)$ is the expected value of C_m.

If the statistical policy is used, we can find the distance l_{thresh} so that $P_m(l_{thresh}) = p_t$. Similarly, c_m can be computed with

$$c_m = 2 \cdot \int_0^{l_{thresh}} P_m(x)dx \qquad (4\text{-}13)$$

Comparing equation (4-12) and (4-13), the coverage with the statistical policy is smaller than that with the deterministic policy. As p_t approaches 0, c_m with the statistical policy approaches that with the deterministic policy. As p_t approaches 1, c_m with the statistical policy approaches 0.

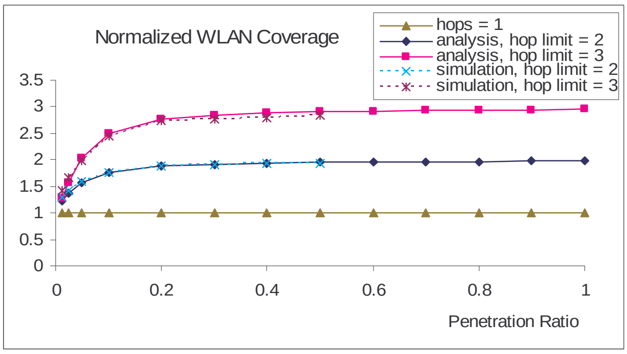


Figure 26: Normalized WLAN coverage for hop limit as 1, 2, and 3 when the deterministic policy is used

Figure 26 shows the coverage of an access point normalized over that of the single-hop case using both analytical models and simulation. In simulation, a WLAN access point is placed in the middle of the studied area along the road. The amount of user data traffic offered to this access point is proportional to its coverage. The 1-hop coverage equals twice the WLAN radio range (or by definition, a normalized value of 1) and remains constant. The analytical results are seen to match very closely the results predicted by the simulation model. This means that the assumptions used in the analytical models (e.g., Poisson point process) are appropriate for the problem addressed here. For multi-hop WLAN, the coverage begins by increasing sharply as the penetration ratio (and thus the density of instrumented vehicles) increases and then almost levels off (close to the maximum coverage). This means that as the instrumented vehicle density increases beyond some saturation value (penetration ratio around 0.3 in this example), there is almost no further extension in coverage. Both simulation and statistical analysis presented here can be used to determine the saturation value of instrumented vehicle density for a given hop limit. Multi-hop forwarding can significantly extend the WLAN coverage even when the penetration ratio is as low as 0.025 (when the average vehicle distance is about 250m).

To interpret this value, remember that the USDOT's ITS Vehicle-Infrastructure Integration (VII) initiative is expecting potentially 10% of the nation's vehicle fleet being instrumented within two years of the commitment to deploy the system. This suggests that significant WLAN coverage extension can be expected in the area studied.

We define the **data traffic load** of an access point as the total amount of user data (not including protocol headers) *scheduled* for the access point under a specific scheduling policy; not all these data eventually arrive, however. Data traffic load reflects user data traffic demand. With our assumption that every vehicle generates the same constant rate of data, the data traffic load of an access point is proportional to the number of vehicles under its coverage. Suppose the instrumented vehicle density is λ and the per vehicle data rate is b. If the deterministic policy is used, the average data traffic load l_{WLAN} of an access point can be computed as

$$l_{WLAN} = \lambda \cdot c_m \cdot b \qquad (4\text{-}14)$$

With the statistical policy, there are missed and false opportunities. Missed opportunities reduce the WLAN data traffic load. While false opportunities do not increase the amount of data traffic eventually arriving at WLAN access points, they may increase the data traffic on forwarding paths when multi-hop forwarding is used. The average data traffic load of an access point can still be computed using equation (4-14) but c_m needs to be computed using equation (4-13). The data traffic load of an access point with the statistical policy is smaller than that with the deterministic policy.

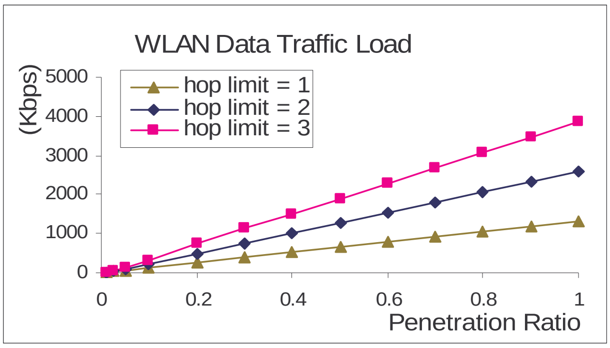


Figure 27: The average data traffic load of a WLAN access point with the deterministic policy

Figure 27 gives some numerical results. The hop limit is set to be 1, 2, and 3, respectively. The data traffic rate per vehicle is 16Kbps, a pretty modest value. Results in [63] [10] show that the user data rate of IEEE802.11b is at most about 5Mbps when UDP is used even though the nominal channel rate is 11Mbps due to various protocol overheads. In Figure 27, the average one-hop data traffic load of a WLAN access point with 100% penetration is about 1.3Mbps, which is well below the IEEE802.11b capacity. However, the average data traffic load when the hop limit is 3 with 100% penetration is about 3.9Mbps, which may exceed the multi-hop IEEE802.11b link capacity due to the reasons that are discussed next.

4.2.6 Usable WLAN Capacity

We define the **usable capacity** of a network system as the maximum effective user data throughput that the system can deliver for user applications to operate normally[7]. It is well known that WLAN channel capacity is reduced by the interference introduced by

[7] When the network is heavily loaded, the loss ratio may be too high to allow normal operation of most user applications. Due to this fact as well as various protocol overheads, most networks systems, e.g., Ethernet or 802.11x, cannot reach its nominal channel rate.

ad-hoc communications and 802.11 MAC fails to achieve the optimum chain schedule [47]. We now shed some light on this issue for the network architectures proposed here.

Figure 28: Illustration of multi-hop interference

Multi-hop forwarding introduces interference that reduces the channel capacity of access points. This can be demonstrated with a simple illustration. In Figure 28, all the data traffic is flowing to the access point in the middle of the figure. Every arrow represents one hop transmission. Similar to [47], we estimate the achievable channel utilization of the WLAN access point. Note that we are modeling the 802.11 DCF unicast transmission preceded by RTS/CTS exchanges and the access point only receives data. If only one hop is allowed, the channel utilization can reach 1 because all the transmissions from either node 1 or 2 are targeting the WLAN AP. If 2-hop forwarding is allowed, when nodes 3 or 4 are transmitting, the access point cannot receive data. This leads to a channel utilization of at most 1/2. If 3-hop forwarding is allowed, in addition to interference due to the 2-hop forwarding, node 1 cannot transmit while node 5 is transmitting, but the access point can receive data from node 2. However this requires a perfect schedule. On average the actual channel utilization should be between 1/2 and 1/3. A transmission more than 3 hops away (e.g., node 7 or 8) should not affect the reception of the access point. Thus a hop limit of more than 3 does not further reduce the channel utilization of the access point.

Furthermore, 802.11 MAC fails to achieve the optimum chain schedule [47] so that some data traffic injected into multi-hop paths is dropped before it reaches the destination. The transmission of these data, however, decreases the network capacity. When the data traffic load increases, the data loss ratio will also increase, eventually becoming too high, severely degrading performance.

We have conducted a series of simulations to study the throughput of a WLAN access point, which is measured as the amount of user data received by the access point per second. A WLAN access point is placed in the middle of the studied area along the freeway. We vary the vehicle data flow rate from 16Kbps to 2048Kbps but fix the vehicle density (penetration ratio = 0.1).

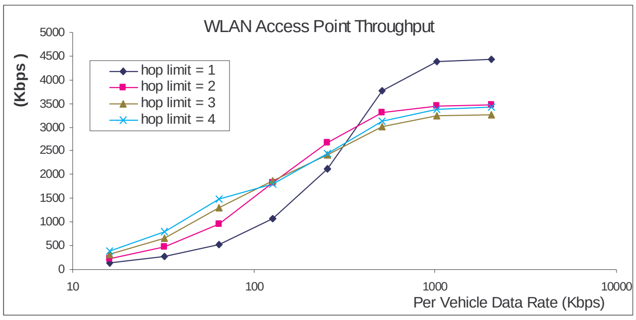


Figure 29: The average WLAN access point throughput as the vehicle data rate increases for different hop limits

As shown in Figure 29, the average throughput of the WLAN access point increases when the data traffic load increases. When the vehicle data flow rate is below approximately 128Kbps, multi-hop forwarding increases the WLAN throughput, e.g., the 2-hop WLAN provides higher throughput than the WLAN last-hop. However, there is a "throughput inversion" when the vehicle data flow rate exceeds approximately 128Kbps, i.e., the throughput degrades when the hop limit increases. This is because the throughput of an access point is affected by both its coverage and its channel capacity. On one hand, multi-hop forwarding extends the coverage area so that more data traffic is allowed to reach this access point. This tends to increase the throughput of the access point. On the other hand, multi-hop forwarding reduces the channel capacity of the access point, potentially reducing its throughput when the data traffic load is high. Furthermore, the impact of these two factors varies over time as vehicles move. For example, it is possible that at a particular moment most vehicles access an access point in one hop even though 3 hops are allowed. But in general, as the data traffic load increases, the limitation on channel capacity becomes more serious. This eventually results in throughput inversion. When the hop limit is greater than or equal to 4, the throughput is not further reduced. This agrees with our analysis. The maximum achievable throughputs are about 4.5, 3.4, 3.2, and 3.4 Mbps when the hop limit is 1, 2, 3, and 4, respectively. These numbers are somewhat different from our channel utilization analysis because hop limit is the maximum number of hops allowed, not the actual hops existing.

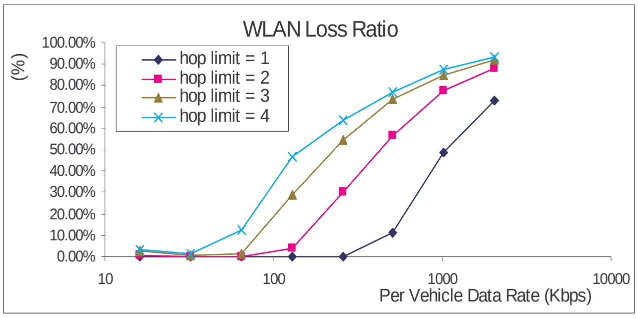


Figure 30: The loss ratio of the WLAN data traffic as the vehicle data flow rate changes for different hop limits

Figure 30 shows the loss ratio as the data rate per vehicle changes. The loss ratio is computed as the percentage of data traffic offered to but not arriving at the access point. A very small loss ratio represents packets in transit. It is not surprising that the loss ratio increases as the data traffic load increases. Examining Figure 29 and Figure 30 together, we find that as the access point throughput approaches its maximum, the loss ratio is already too high to render the system useful (except when the hop limit is 1). The WLAN last-hop approach could achieve the useful capacity close to the theoretical limit. If multi-hop forwarding is allowed, the actual usable capacity is much lower than the maximum achievable throughput. To have the loss ratio below 20%, for example, the throughputs are about 4.2, 2.3, 1.7, 1.6 Mbps when the hop limit is 1, 2, 3, and 4, respectively.

The above results indicate that 802.11b systems may not provide sufficient capacity under certain circumstances, e.g., high data rate per vehicle, high instrumented density, or multi-hop forwarding. Under these circumstances, last-hop may be preferred over multi-hop. It may also be necessary to deploy WLAN systems with higher channel capacity (e.g., 802.11a/g).

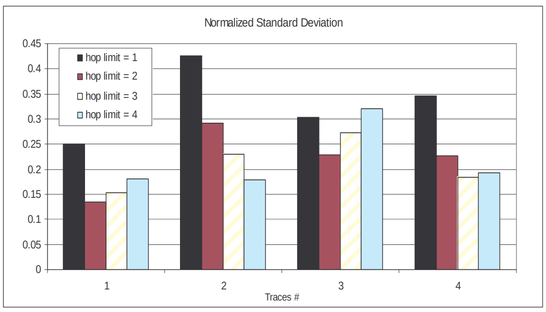
Downstream data traffic may exhibit different throughput characteristics as all the downstream data traffic originates from access points and receivers are vehicles. Further exploration of this issue is an area of future research.

4.2.7 Impact of Vehicle Traffic Dynamics

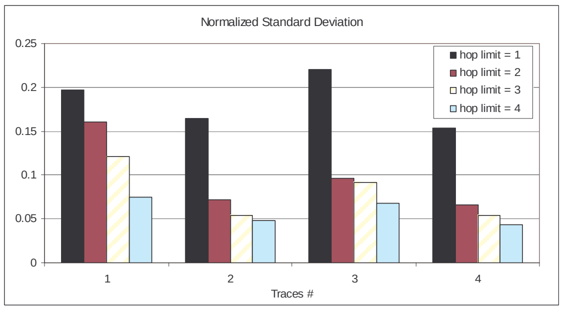
Because vehicle traffic is moving constantly and traffic conditions change over time, intuitively the vehicle traffic dynamics will have significant impact on communication system design. Extra resources (e.g., link bandwidth, data buffer) and/or special designs (e.g., queue sharing algorithms) are often required to accommodate surges in data traffic caused by instrumented vehicle traffic increase. In this section, we will give an example showing the impact of instrumented vehicle traffic dynamics. Instrumented vehicle traffic dynamics are caused by both vehicle traffic and instrumentation dynamics.

Vehicle traffic conditions change drastically throughout a day. Here we examine the variance of the number of instrumented vehicles connected to WLAN during a period of time when the vehicle conditions are considered relatively stable. In our study, we examine the morning rush hours along the selected I-75 corridor. The variance of instrumented vehicles within the coverage of a WLAN access point will cause the variance of data traffic load for this access point. We try to identify the factors that affect the variance.

With our assumption that every instrumented vehicle generates the same constant rate of data, the data traffic load of an access point is proportional to the number of instrumented vehicles under its coverage. We collect the second-by-second data traffic load of an access point through simulation. We use the standard deviation normalized over the average data traffic load to capture the *scale* of variance.



(a) Low Instrumented Vehicle Density (Penetration Ratio = 0.1)



(b) High Instrumented Vehicle Density (Penetration Ratio = 0.5)

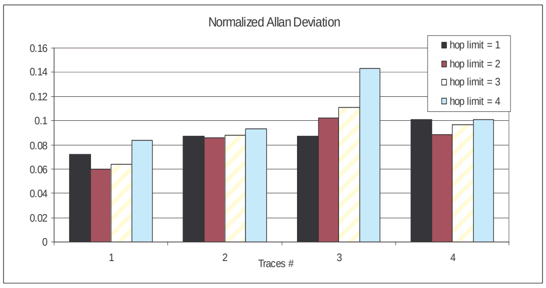
Figure 31: Normalized standard deviation of the number of instrumented vehicles connected to WLAN over its mean for a time scale of 1 second for different hop limits

Figure 31 shows the normalized standard deviation for both high and low instrumented vehicle densities for 4 vehicle traffic traces. Penetration ratio of 0.1 (0.5) represents low (high) instrumented vehicle density. There are several observations. First, the variance can not be neglected. It can reach more than 0.4 for low instrumented vehicle density and more than 0.2 for high instrumented vehicle density. This shows that during a period of relatively stable vehicle traffic, it may require additional resources to accommodate surges of data traffic. Second, the variance when the instrumented vehicle density is low is higher on average than that when the instrumented vehicle density is high. Third, when the instrumented traffic density is low, we do not find correlations between the variance of the number of instrumented vehicles connected to the WLAN and the hop limit. However, when the instrumented traffic density is high, higher hop limit consistently yields lower variance.

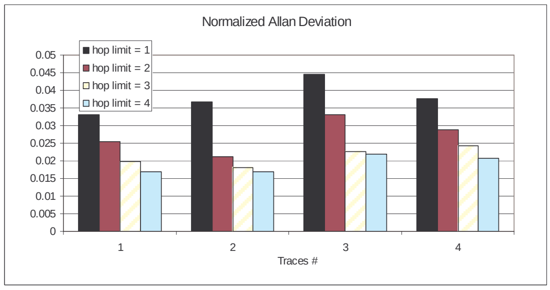
We can explain the above observations by considering two factors. First, when the number of instrumented vehicles connected to the WLAN increases, this number tends to be more stable since the base is becoming larger (large number theory). This explains the correlation between the variance and the instrumented vehicle density. Second, due to high vehicle mobility, long forwarding paths (in the number of hops) are subject to frequent breakage/formation, resulting in greater variance of instrumented vehicles connected to the WLAN. But the impact of the second factor diminishes as the instrumented vehicle density increases (leading to more stable paths). When the instrumented vehicle density is low, more hops means more instrumented vehicles are connected to the WLAN, but the path is also easier to vary. The combined effect of these two factors results in no correlation between the variance of the number of instrumented vehicles connected to the WLAN and the hop limit. When the instrumented vehicle density is high, the second factor almost has no impact. More hops means more vehicles connected to the WLAN and thus the number of instrumented vehicles connected to the WLAN tends to be more stable.

Another important property of time variance is the *rate* of change, which characterizes how quickly the quantity of interest changes. For example, the number of instrumented vehicles connected to a WLAN access point may increase by 10% every minute or every hour. These two cases may require different network resource allocation strategies. One way to capture the speed of variance is to plot the Allan deviation. Allan deviation differs from the standard deviation in that it uses the difference between subsequent samples in time rather than between samples and the mean as standard deviation does. The formula to compute the Allan deviation of a data sets s_i is:

$$\text{Allan deviation} = \sqrt{\frac{1}{2 \cdot n} \sum_{i=2}^{n} (s_i - s_{i-1})^2} \qquad (4\text{-}15)$$



(a) Low Instrumented Vehicle Density (Penetration Ratio = 0.1)



(b) High Instrumented Vehicle Density (Penetration Ratio = 0.5)

Figure 32: Normalized Allan deviation of the number of vehicles connected to a WLAN access point over its mean for a time scale of 1 second for different hop limits

Figure 32 shows the normalized Allan deviation for both high and low instrumented vehicle densities for 4 vehicle traffic traces. Figure 32 is similar to Figure 31. We can therefore derive similar observations as above. In addition, the rate of change is slow when the instrumented vehicle density is high. When the instrumented vehicle density is high, the number of instrumented vehicles connected to the WLAN only varies by a few percent from one second to the next, one implication of which is that the average during several seconds can be a useful predictor of near-future value. This fact can be exploited to design adaptation strategies.

4.2.8 Number of WLAN Access Points

We now discuss how many WLAN access points are required to provide continuous coverage for 3 architectures: WLAN last-hop, WWAN + WLAN last-hop, and WWAN + multi-hop WLAN. Due to the rapid growth of the WLAN bandwidth, we assume the WLAN capacity is not a limiting factor. Normally the WWAN capacity is fixed and a limiting factor. It is also assumed that WLAN access points are placed at equal distances between neighbors. One access point is placed on each end of the road so that a vehicle is always driving between two access points. In the following analysis, we assume the deterministic policy is used. Similar analysis can also be done if the statistical policy is assumed.

Suppose the length of the studied road segment is l, the average number of instrumented vehicles on the road segment is n, the WLAN radio range is r, the per vehicle data rate is b, the aggregated WWAN data rate for this area is B, and the number of access points is k.

For the WLAN last-hop architecture, it is easy to see that

$$k \geq l/2r + 1 \tag{4-16}$$

For a WWAN+WLAN hybrid architecture, the WWAN data traffic load should not exceed the WWAN capacity. Hence, for the WWAN + WLAN last-hop architecture, we have

$$(1-\frac{(k-1)\cdot 2\cdot r}{l})\cdot n\cdot b < B \Rightarrow k > (1-\frac{B}{n\cdot b})\frac{l}{2\cdot r}+1 \qquad (4\text{-}17)$$

For the WWAN + m-hop WLAN architecture, we use an iterative algorithm that is illustrated with the pseudo-code in Algorithm 3.

```
(1) k=k1
(2) while(true)
(3)     d = l/(k-1)
(4)     if((1-Pd,m)*n*b > B)
(5)             output(k+1)
(6)             exit
(7)     else
(8)             k--
```

Algorithm 3

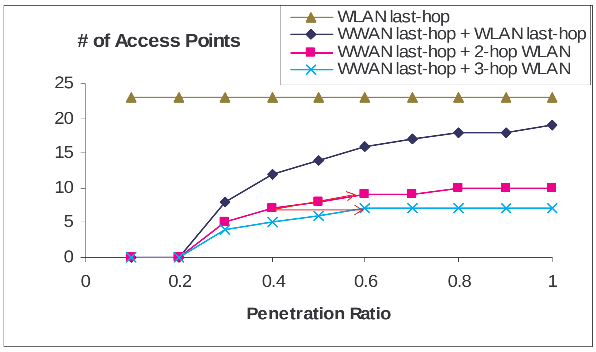


Figure 33: Number of access points required for WLAN last-hop and hybrid WWAN+WLAN architectures with hop limit equals 1, 2, and 3

In the Algorithm 3, k is assigned an initial value k_1 (line 1), i.e., the number of access points required for WWAN + WLAN last-hop, and then decrements until the WWAN traffic load is more than the WWAN capacity. When the algorithm terminates (line 6), k+1 is output as the result.

Figure 33 shows a concrete example. The numerical values of system parameters are: l = 11,000 m, n = 1800 * penetration ratio, r = 250 m, b = 16 Kbps, B = 6 Mbps[8]. With a WLAN last-hop architecture, a fixed number of access points are required (23 in this

[8] We assume 3 WWAN cells are covering the studied area, each with a link of 2 Mbps.

example) regardless of vehicle traffic condition. The hybrid architecture can significantly reduce the number of required WLAN access points because it is not necessary to deploy WLAN access points to cover everywhere. Multi-hop forwarding (even one more hop) can help reduce this number even further. This can lead to a reduction in construction and maintenance costs.

Here we also show an adaptation approach to accommodate varying instrumented vehicle traffic. In a hybrid architecture, either the number of access points or the hop limit can be adjusted for adaptation. For example, suppose the penetration ratio is 0.4, there are 7 access points, and the hop limit is 2. Now because vehicle traffic condition changes, the instrumented vehicle density changes as if the penetration ratio were increased to 0.6. We can either notify vehicles to change the hop limit to 3, or increase the number of access points to 9, as illustrated by the two state transitions in Figure 33. The latter can be realized by activating some backup systems or deploying mobile access points.

The number of access points computed here only considers the average behavior. If system dynamics are taken into account, more access points are required.

4.2.9 Data Throughput Per Vehicle

The above evaluation is conducted from the perspective of infrastructure deployment. We now examine the data throughput experienced by individual vehicles under various network architectures. The vehicle data throughput is defined as the amount of user data from this vehicle arriving at the infrastructure every second.

We first briefly describe how to estimate data throughput per vehicle using analytical models. In a hybrid WWAN + WLAN architecture, we can estimate the expected WWAN throughput that a vehicle may experience when moving in an area not covered by the WLAN. Suppose the average number of instrumented vehicles on the studied area is n, the aggregated WWAN bandwidth for this area is B, WLAN access points are separated by a distance of d, the hop limit is m, and b_w is the vehicle's expected WWAN throughput, it is easy to see that

$$b_w = \frac{B}{n \cdot (1 - P_{d,m})} \tag{4-18}$$

The above formula indicates that increasing $P_{d,m}$ (adding more WLAN access points or increasing the hop limit) can improve b_w. This is not surprising since there are fewer vehicles sharing the WWAN when $P_{d,m}$ is higher.

As a vehicle is moving, its instantaneous throughput varies depending on its connectivity and the channel contention. With equation (4-18) and the WLAN throughput analysis such as [49], we can estimate the average throughput a vehicle may experience when using the WLAN and WWAN, respectively.

Simulations were performed to examine the average vehicle throughput for three architectures: WWAN last-hop, WWAN last-hop + WLAN last-hop, and WWAN last-hop + multi-hop WLAN. We use 3 WWAN cells, each with a 2 Mbps link, to cover the entire studied road with each one covering about 1/3 of the road segment. Seven IEEE802.11b access points are placed along the road with one on each end of the road. This number is chosen so that the 3-hop coverage of each access point is not overlapped with each other. We vary the vehicle data flow rate from 16Kbps to 512Kbps but fix the instrumented vehicle density (penetration ratio = 0.1).

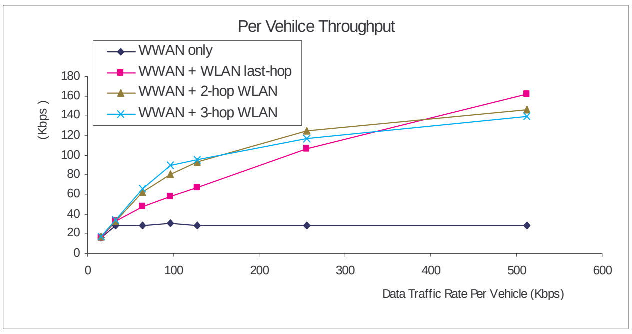


Figure 34: The average vehicle throughput as the function of the vehicle data flow rate for different hop limits

As shown in Figure 34, the WWAN last-hop architecture can only provide an average per-vehicle data throughput up to only 30Kbps. With the addition of the WLAN, the vehicle throughput can reach more than 100Kbps. We also see the "throughput inversion" approximately around the vehicle data flow rate of 128Kbps due to the WLAN capacity limit as discussed in Chapter 4.2.6.

4.3 Implication

Designing a vehicular network requires a careful assessment of tradeoffs between cost and performance. The evaluation described above gives some insights into the main wireless network infrastructure building blocks and their interactions. Different network architectures are appropriate in different environments. In addition, the design of a vehicular network must address **changing vehicle traffic conditions over time**. Next we will discuss the application of this research in vehicular network infrastructure design.

The inputs to the architecture design process include the existing network facilities, anticipated user data traffic and coverage demands, network usage pattern, vehicle traffic conditions, and projected penetration ratio. The output of the design process includes the configuration of the network, including the type of the WWAN and its capacity, type of the

WLAN, number of WLAN access points, hop limit, and adaptation strategies. There are usually many feasible network configurations.

One can envision two types of connectivity requirements: continuous connectivity or intermittent connectivity [63].

WWAN can provide continuous connectivity but does not scale well to support a large number of instrumented vehicles (Chapter 4.2.4). WWAN alone is only a feasible design option when bandwidth requirements are limited.

An alternative is to place a sufficient number of WLAN access points along the road to provide continuous last-hop access. This design is attractive due to its simplicity and easy deployment and can provide high data throughput to travelers. The major drawback of this architecture is the high cost to install and maintain a large number of roadside access points (at least one every 1/3 mile if the WLAN radio range is assumed 250m). In places where the cost is not a major concern, this may be a good choice.

In places whether the WWAN service already exists but does not offer adequate capacity, WLAN access points can be placed along the road to satisfy the system capacity requirement. Travelers might experience higher throughput within the coverage of the WLAN but lower throughput otherwise (Chapter 4.2.9). Furthermore such a hybrid architecture requires much less number of WLAN access points than WLAN last-hop architecture while still maintaining continuous connectivity (see Chapter 4.2.8). This number can be further reduced by utilizing multi-hop forwarding. Allowing multi-hop forwarding has both positive and negative features. As shown by our evaluation, the major benefit of multi-hop WLAN is the increased coverage of each WLAN access point (we observed significant coverage extension even when the average vehicle distance is about the radio range in Chapter 4.2.5), resulting in reduced number of access points required to offload enough data traffic from the WWAN. Major drawbacks with this approach include reduced channel capacity due to interference (see Chapter 4.2.6) and increased system complexity in routing, accounting and security. The vulnerability of multi-hop paths has already been mitigated by the WWAN. Therefore data traffic load, instrumented vehicle density and other factors (e.g., delay & routing overhead) will determine the hop limit. Data traffic load dictates whether the reduced capacity caused by multi-hop forwarding becomes a limiting factor. Instrumented vehicle density decides whether there are significant benefits in terms of coverage extension. Due to the fast-growth of WLAN capacity, the WLAN channel capacity should not be a major concern in most cases. The major issue then becomes whether the reduced cost is sufficient to justify the additional system complexity. Use of multi-hop forwarding seems questionable due to the additional complexity it entails, except in places where there are limitations on the number of access points that can be installed, e.g., due to cost or terrain constraints. However, even if multi-hop forwarding is disabled in general, it is still beneficial for vehicles to cooperate voluntarily at times in order to improve performance in areas of poor wireless coverage (see Chapter 5).

If intermittent connectivity is deemed sufficient, i.e. the intended applications can tolerate some loss of connectivity [63], a WLAN-based architecture is preferred because it allows quick deployment and relatively low construction cost. The intermittence requirement determines the density of access points that must be deployed. Again the decision to use multi-hop forwarding is governed by a tradeoff between cost and system complexity. For every combination of the vehicle traffic condition, WLAN access point placement and hop limit, the connectivity probability in every location can be estimated (Chapter 4.2.2). If only last-hop is allowed, the probability is either 1 or 0 in a particular location. If multi-hop forwarding is allowed, the probability could be some intermediate value.

Vehicle traffic conditions vary continuously over time. The variance often cannot be neglected in designing communication systems even during a period of relatively stable vehicle traffic condition (Chapter 4.2.7). Precise, fine-grained variations in vehicle traffic are difficult or impossible to predict. Some coarse-grained vehicle traffic variation can be predicted, e.g., due to rush hours or road construction, while some can not, e.g., due to accidents. System design should take this variance into consideration. One solution is to overprovision the system, i.e. allocate additional resources (e.g., link capacity, processing power and data buffers) to accommodate surges in user demands. An alternative is to design adaptation strategies: the vehicle traffic load is monitored continuously and the system configuration is changed accordingly, e.g., changing hop limit, activating/deactivating backup systems, or deploying mobile systems (Chapter 4.2.8). The overprovision approach can be applied to handle variations that are difficult to predict while the adaptation approach is effective in dealing with predicable variations. These two approaches can be combined in practice.

4.4 Conclusion

In this chapter, we identified several design options for infrastructure-based vehicular networks and studied the tradeoffs offered by various wireless technologies in a realistic vehicular environment. We developed analytical models to provide some insights and derive useful settings for detailed evaluation. Simulation is employed to verify the statistical analysis and examine additional issues that are difficult or impossible to capture with analytical models. The evaluation results lead to some insights concerning issues such as the use of multi-hop forwarding in vehicular networks. Most of the concepts involved are intuitive; our contributions lie on showing their significance in the context of vehicular networks and reveal some issues not so straightforward.

The numerical results presented in this chapter are limited by the assumptions concerning data traffic pattern, communication parameters and the vehicle traffic model that was used (however, we do expect this model represents a typical urban traffic setting.) Similar studies should be conducted for a broad range of settings (e.g., upstream data traffic, mix of upstream and downstream traffic, self-adaptive data traffic, different communication parameters, vehicle traffic conditions in different time and areas, etc.) However, our evaluation methodology (including both simulation and analytical models) is sufficiently general to be useful for further studies.

We evaluated several architectures primarily from the perspective of system deployment. The architectural implications on protocol and application design need further exploration. For example, although the WWAN+WLAN hybrid architecture shows many benefits, it also increases the system complexity in several ways: routing, handoff management, billing and accounting (especially when multi-hop forwarding is involved.) In particular, techniques such as data striping [67] and vertical handoffs [77] in multi-radio systems as well as mobility management [39] have direct relevance.

CHAPTER

5 MEASURING SHORT RANGE COMMUNICATIONS

5.1 Introduction

An in-depth understanding of communication characteristics in a vehicular environment is needed to provide the groundwork for realizing reliable mobile communication services. Also simulation models need to be verified with field data. Vehicular communications are distinct from other types of wireless communication because of the high mobility of vehicles and the environment in which they operate. As described in 1.5.2, the initial works in measuring the performance of short-range communications have not yet related the road environment to communication performance.

In this chapter, we focus on identifying factors in the road environment that affect wireless communication performance [91]. Instead of trying to cover many areas as in [74], our experiments were concentrated on a specific highway environment, thereby allowing a deeper understanding of factors affecting communication performance. Single hop experiments were used to characterize V2V and V2R communications and identify the factors that affect communication performance. We then demonstrated how multi-hop communications can help overcome the obstacles presented by the physical environment.

5.2 Experimental Design

We have performed three different types of experiments: vehicle to roadside station communication, vehicle to vehicle communication, and multi-hop communication. For these experiments, we used a laptop computer in each vehicle running Red Hat Linux 9, an

ORiNOCO 802.11b gold card with a omni-directional external antenna placed on the roof of the vehicle, and a Garmin 72 GPS receiver (with WAAS correction). The external antenna is the Proxim ORiNOCO Range Extender Antenna (010096/E), 2400-2500 MHz. 2.5 dBi, 0-55C, with a proprietary antenna connector. This antenna is attached to the ORiNOCO 802.11b Gold card (via the connector). The antenna is designed for indoor use. We modified the antenna by removing the rubber base, enlarging the screw hole, and attaching the antenna to the magnetic base mount of the Magnetic-Mount CB Antenna sold by Radio Shack (Catalog Number 21-989 at www.radioshack.com.) Because it is not waterproof, we only used it in good weather. We have since designed a waterproof plexiglass housing and are now performing tests during rain conditions. Since the standard DSRC hardware is not commercially available at this time, we used 802.11b wireless interfaces instead. The radio transceivers were configured to work in IBSS ad-hoc mode with power management turned off and a fixed data rate. The current IBSS implementation in the firmware of ORiNOCO 802.11b gold card has certain bugs that result in the change of cell id and/or the communication channel when a card is alone. To circumvent this problem, we always doubled the number of cards used so that a network card was never alone. This solves the problem most of the time. The GPS reports latitude, longitude, speed, and heading of the vehicle every 2 seconds[9]. The location information reported by GPS has an accuracy of 5-7 meters. We obtained location information every second through interpolation.

We used IPerf [9] in conjunction with GPS readings to measure network performance measurement. All communications were conducted using broadcast at

[9] The vehicle drives for about 60m for 2 seconds.

2Mbps that allows us to explore raw communication performance because there is no RTS/CTS/ACK and retransmissions. For multi-hop communications, we also implemented a Linux kernel module package for data forwarding. All the participants logged time, location, velocity (from GPS), data packets received/sent, and signal quality. All logs were associated and merged together to obtain the results presented. Capturing vehicle traffic condition is also important, but was not implicitly done for these experiments, although experiments were conducted under non-congested conditions. Characterization of dynamic vehicle traffic conditions in a way that is meaningful to wireless communication is still a research topic requiring future exploration.

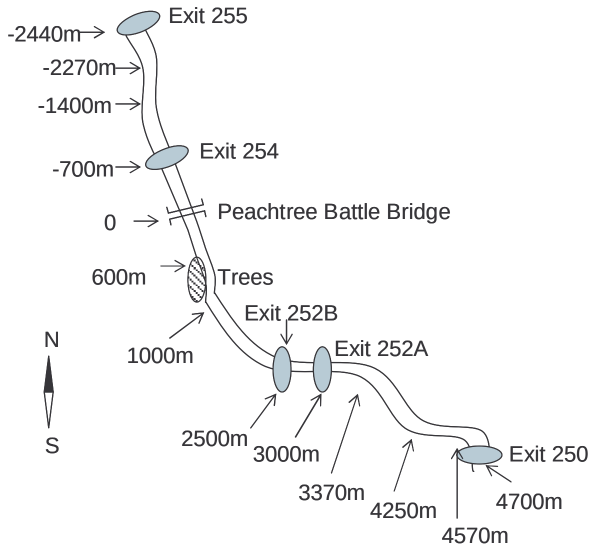


Figure 35: Road illustration (Not to Scale)

The experiments were conducted in the northwest sector of Atlanta, GA along I-75 between Exit 250 (14th Street) and Exit 255 (West Paces Ferry) as shown in Figure 35. The distance shown in the figure is the direct distance to Peachtree Battle Bridge, an overpass between Exit 252B (Howell Mill Road) and Exit 254 (Moors Mill Road) of I-75. Negative distances represent the distance to the north of Peachtree Battle Bridge, while positive distances correspond to the south. This section of the I-75 corridor has five regular lanes in addition to one HOV lane in most parts of the study area. Vehicle traffic in opposite directions is separated by a concrete median barrier. Similar configurations can be found in many major U.S. urban areas. All the experiments were conducted between 2pm and 5pm under non-congested and non-inclement weather conditions over a period of half a year. As our focus is to measure the impact of road environment on wireless communication, we fixed most of the communication parameters during the experiments. The experiment for each system configuration was repeated at least 5 times.

5.3 Vehicle-to-Roadside Communications

We first measured the wireless communication performance between a fixed roadside station and a moving vehicle. The fixed roadside station was also equipped with the same equipment as the in-vehicle system. The roadside station was placed above the median on Peachtree Battle Bridge. An instrumented vehicle traveled in loops between Exit 252B and 255 along I-75. The moving vehicle constantly broadcasted packets of 1470 bytes. The data packets were sent as rapidly as possible at an average rate of about 150 packets/s. The vehicle was traveling in the rightmost lane whenever possible. We have experimented with different configurations of the roadside station: in separate experiments it was situated in either the northern (facing Exit 254) or southern (facing Exit 252B) side of the bridge, and the external antenna was either placed on a table (0.7m) or on a much higher tripod (1.8m). The results presented in this section are the aggregation of 5 laps. The distance is plotted the same way as in Figure 35, with the origin located at Peachtree Battle Bridge. The *success ratio* is defined as the fraction of packets transmitted by the vehicle that are successfully received by the roadside station. Each data point represents a distance scale of 30m.

5.3.1 Without vs. With External Antenna

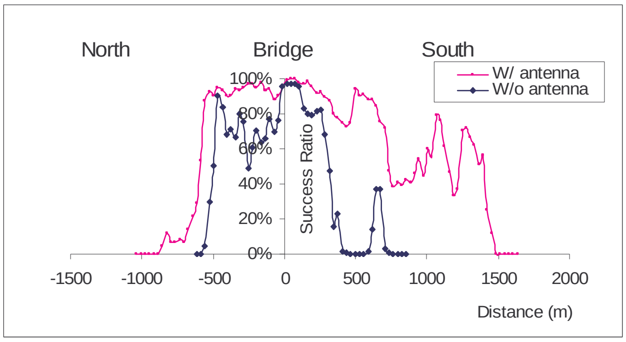


Figure 36: The communication performance over the distance between the vehicle and roadside station w/ and w/o an antenna on the vehicle

In Figure 36 we compare the performance with and without an external antenna mounted on the roof of the vehicle. The roadside station was in the southern side of the bridge with the external antenna on the table. The vehicle was driving south (from Exit 255 to Exit 252B). Communication is much worse without the antenna. We attribute this to the obstruction of the vehicle body. In all experiments, the laptop was placed on the passenger side front seat of the vehicle. When the vehicle is driving toward the roadside station (from Exit 255 to Peachtree Battle Bridge), the impact is less severe because the signal only needs to penetrate the windshield. When the vehicle is driving away from the roadside station (from Peachtree Battle Bridge to Exit 252B), the impact is more severe because the signal is obstructed by the entire vehicle body. While placing the antenna on the roof of the vehicle may amplify the signal somewhat (there is attenuation in the cable and connector), it eliminates much of the impact of the vehicle body, which appears to be a dominant factor in improving performance. To mitigate the uncertainties brought about by the vehicle body, all remaining experiments placed an external antenna on the roof of the vehicle.

5.3.2 A Detailed Scenario

We now examine the V2R communication in detail by using one configuration of the roadside station. The roadside station was situated in the southern side of the bridge and the antenna placed on a table. We monitored the noise level of the roadside station and found it remained constant at approximately -100dbm.

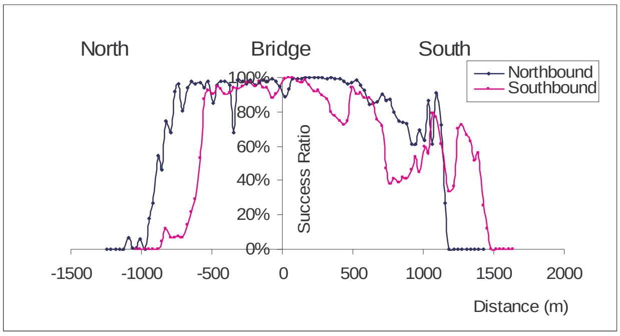


Figure 37: The communication performance over the distance between a vehicle and the roadside station as the vehicle moved in both directions

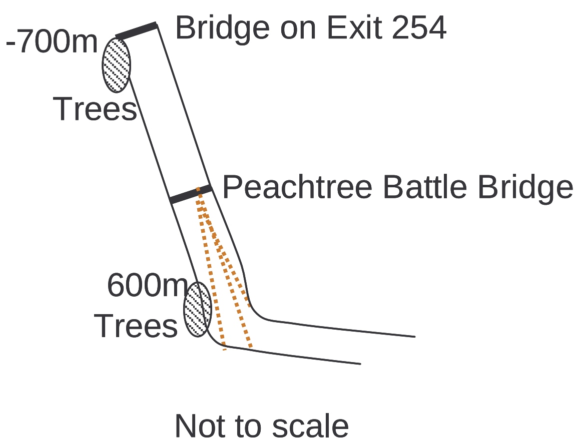


Figure 38: Illustration of road geometry

Figure 37 illustrates the communication performance when the vehicle traveling in the northbound and southbound directions, respectively. First, we have observed intermediate loss rates in some areas. This means there is no clear cut here on whether communication is possible or not, suggesting a need to map communication performance. Aguayo et al. [8] obtained similar observations in a mesh network. One can derive effective communication range corresponding to different success ratios from this figure. For example, more than 300m of effective communication range with a success ratio over 80% is observed on both sides of the roadside station (a success ratio of 80% means that a message can be received with 96% probability if transmitted twice). Secondly, the communication performance depends largely on the geometry of the road, obstructions, and vehicle locations in addition to the signal attenuation over distance. Consider the performance with the vehicle traveling south. Figure 38 illustrates the road geometry around Peachtree Battle Bridge. To the north of Peachtree Battle Bridge, the performance remains good until a sharp drop around -600m. This is approximately the location of a group of trees, immediately followed by a bridge crossing the highway on Exit 254. To the south of Peachtree Battle Bridge, the performance degrades over distance in general. There is also a significant performance dip around the distance of about 850m. This can also be accounted for by blockage due to the tops of trees. After this point there is an improvement in performance for a short length as the vehicle resumes Line Of Sight (LOS) with the roadside station. After that, the communication is lost due to a curve in the road. When we compare northbound and southbound performance, we can see the impact of vehicle location. To the north of Peachtree Battle Bridge, good communication range with northbound vehicles extends until Exit 254 (about -750m), which is further than that with southbound vehicles because northbound vehicles' view is clear until reaching the bridge over Exit 254, while the southbound vehicles' view is shorter due to trees before the bridge.

For the same reason, to the south of Peachtree Battle Bridge, the communication range is longer for vehicles traveling south compared to vehicles traveling north (see Figure 38). During one pass-through, a vehicle can send about 10MB of broadcast data to the road side station.

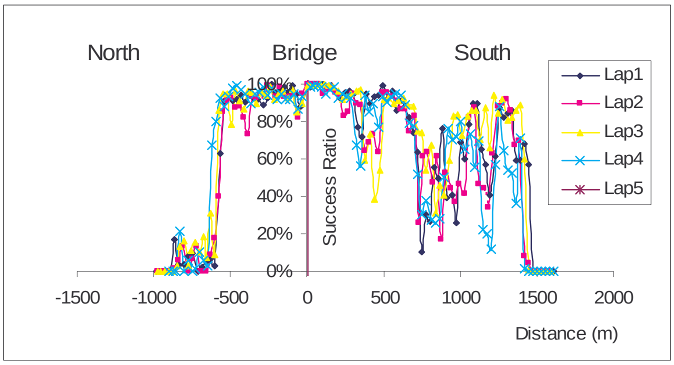


Figure 39: Lap-by-lap results

Figure 39 presents the lap-by-lap measurements as the vehicle traveling south. It can be seen that performance over different laps remains consistent within the range allowed by the GPS sample rate (2 seconds) and GPS reading accuracy (5-7m). This suggests that the performance over location is fairly predictable when one communication endpoint is a fixed roadside station.

From the above results we observe that in addition to the signal attenuation over distance, the factors that affect LOS between the moving vehicle and the roadside station dominate in affecting communication performance. These factors include the geometry of the road (e.g., curves), obstructions (e.g., trees) and the location of the vehicle (e.g., lane occupancy).

5.3.3 Roadside Station Configurations

We now examine the impact of different configurations of the roadside station.

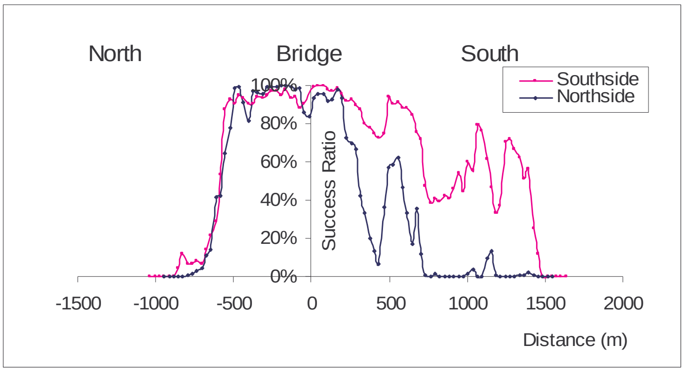


Figure 40: The impact of the placement of the roadside station when the vehicle was driving south and the antenna was placed on a table

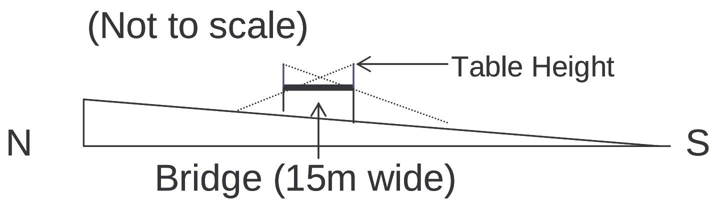
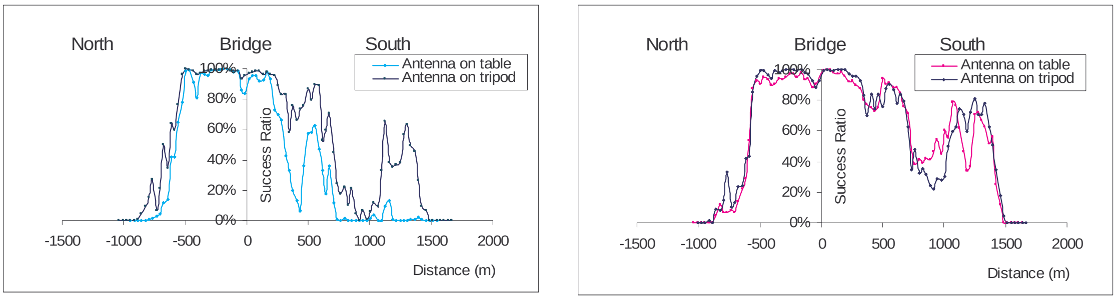


Figure 41: Side view of the road (antenna on the table)

We first placed the roadside station in different sides of Peachtree Battle Bridge. This is to test the effect of the bridge body (about 15m) on the communication. As illustrated in Figure 40, the body of the bridge almost has no impact on the communication performance to the north of the bridge. However, the body of the bridge significantly affects the communication to the south of the bridge. This is due to the elevation of the road. In the area studied, the elevation of the road gradually declines from the north to the south as shown in Figure 41. Therefore, when the vehicle is traveling in the north of the bridge (in either direction), the roadside station can almost always maintain LOS with the vehicle even when the roadside station is placed on the southern side of the bridge. However, when the vehicle is traveling in the south of the bridge, the bridge significantly blocks the signal when the roadside station is placed on the northern side of the bridge.



(a) roadside station in the northern side of the bridge (b) roadside station in the southern side of the bridge

Figure 42: The impact of the height of the antenna for the roadside station when the vehicle was moving south

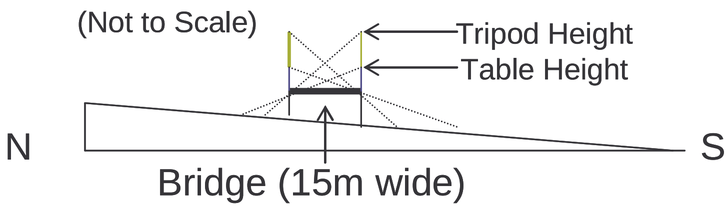


Figure 43: Side view of the road (antenna on the tripod)

We next experimented by raising the height of the roadside station antenna by placing it on a tripod (1.8m). As shown in Figure 42 (a), when the roadside station is in the northern side of the bridge, raising the antenna height significantly improves the reception to the south of the bridge due to the increased area with LOS as shown in Figure 43. However, as demonstrated in Figure 42 (b), when the roadside station is in the southern side of the bridge, raising the antenna only has minor influence on the communication because it does not significantly improve LOS.

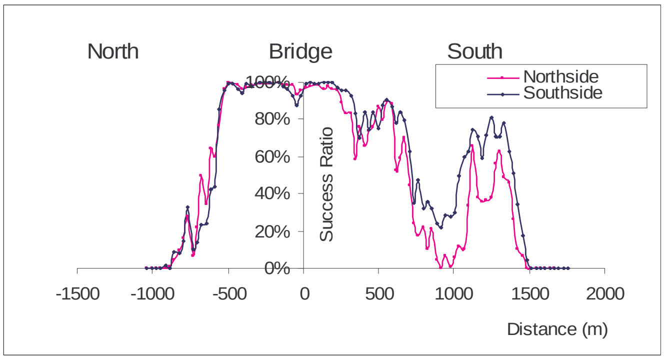


Figure 44: The impact of the placement of the roadside station when the vehicle moved south and the antenna was on a tripod

Figure 44 shows that the placement of the roadside station only has minor impact on communication performance once the antenna is on a tripod. Comparing Figure 44 and Figure 40, it indicates that the bridge body is no longer a significant factor once the antenna is elevated. This can be easily explained by Figure 43.

The above results emphasize the importance of properly locating the roadside station to maximize communication performance. The antenna should be placed in a way to maintain good LOS with vehicles.

5.3.4 Discussion

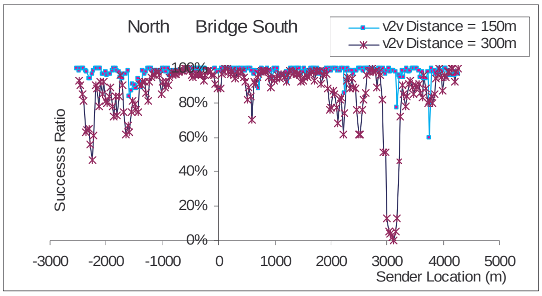
In V2R communications, distance and LOS are two major factors affecting communication performance. Road curves and obstacles, such as trees and bridges, can obstruct the signal, severely degrading communication and leaving “dead spots”. Providing complete coverage may require the deployment of a large number of roadside access points due to the limited Wi-fi coverage. Another approach is to leverage multi-hop V2V communications so that a vehicle can reach the roadside station even if it does not have direct access to it. This approach will be examined later. Another important observation is that communication performance is relatively predictable. One consequence is that some adaptation techniques (e.g., varying packet size [46]) may work well with communication profiles. With our experiment setup, the surrounding vehicle traffic (e.g., a truck in front) has little impact on communication since the roadside station was elevated.

5.4 Vehicle-to-Vehicle Communications

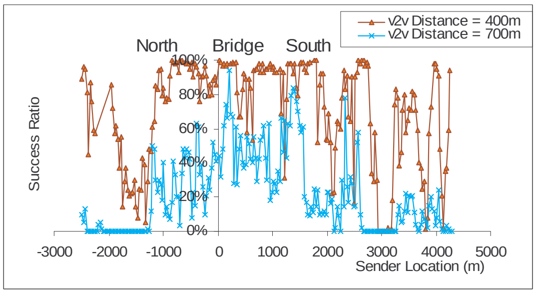
The next set of experiments measured communication performance between two vehicles. We first measured two vehicles following each other in the same lane. We then performed a number of tests of two vehicles crossing each other while driving in opposite directions.

5.4.1 Car Following Scenario

These experiments were conducted on I-75 between Exit 250 and Exit 255 (see Figure 35). To study the effect of distance on performance we varied the space between the two vehicles for different laps. The trailing vehicle was always the sender in these experiments, constantly broadcasting packets. The two vehicles remained in the second lane from the right if possible. The performance plots are generated by aggregating results from 10 laps.



(a) Short inter-vehicle distance



(b) Long inter-vehicle distance

Figure 45: The communication performance between two vehicles over sender locations when vehicles were moving south for different inter-vehicle distances

First, let us examine southbound performance. In Figure 45, the sender location is presented as the distance to Peachtree Battle Bridge. The reception is nearly perfect when the distance between the two vehicles is about 150m. In this case, there is typically one car between the two vehicles. Once the inter-vehicle distance increases to about 300m, the communication performance is good in general, except at some points on the road. By associating the communication performance with the road geometry (Figure 35), we can identify the causes of these dips in performance. Without exception, performance dips all correspond to road curves or overpass bridges. For example, the dips at distance "3000m" and "2500m" can be explained by the bridges on Exit 252A and Exit 252B, respectively. The dips at distance "600m" and "-1500m" occur at points of roadway curves. When the inter-vehicle distance reaches 400m, good communication is only possible during spots between overpass bridges and where the road is relatively straight. V2V communication performance is very poor once the inter-vehicle distance reaches 700m.

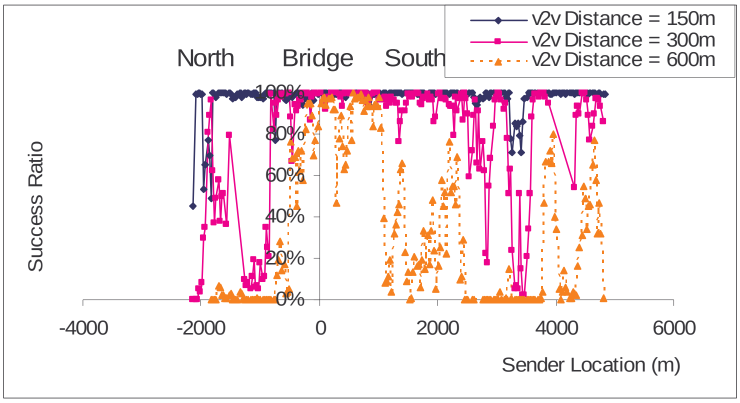


Figure 46: The communication performance between two vehicles over sender locations when vehicles were moving north

Similar observations as above can be made when we examine communication performance when vehicles are driving north (Figure 46). The reception is generally very good when the two vehicles are within 150m. When the inter-vehicle distance increases to about 300m, the communication performance is good in general except for some dips corresponding to roadway curves and bridges. The inter-vehicle communication is almost impossible once the distance between vehicles reaches 600m.

The above results demonstrate that V2V communication performance depends largely on inter-vehicle distance. The communication performance is generally good when two vehicles are within 300m. On the other hand, the effective communication range can extend to as far as 600m in some areas where the road is relatively straight and free of major obstructions.

5.4.2 Car Crossing Scenario

The experiments were conducted on I-75 between Exit 252B and Exit 254. The two vehicles always stayed in the rightmost lane, and were separated by 8 lanes of traffic and a concrete median. Average vehicle speed was around 100 km/h. The chosen area is relatively straight and amenable to wireless communication. Other settings were similar to those described before.

Figure 47 shows the effective communication range of all the measurements. Here the effective communication range is defined as the maximum distance between 2 vehicles when the success ratio is over 80%. Most test cases have more than 200m of effective communication range (the longest range can reach around 1000m). The average time for effective communication is about 21 seconds. This shows that a significant amount of data can be passed from one vehicle to another when they cross each other in this scenario (about 4.6MB for 21 seconds).

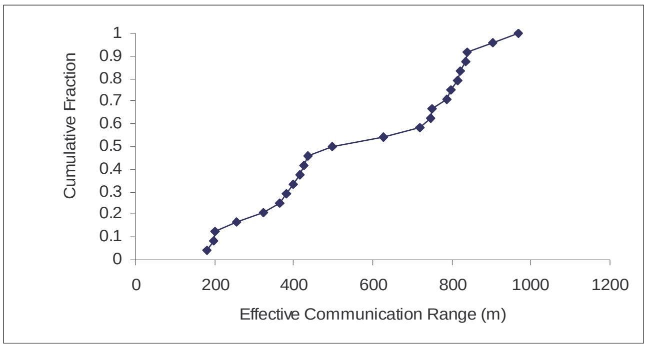


Figure 47: The cumulative distribution of effective communication ranges between 2 vehicles moving in opposite directions

5.5 Multi-hop Communication Experiments

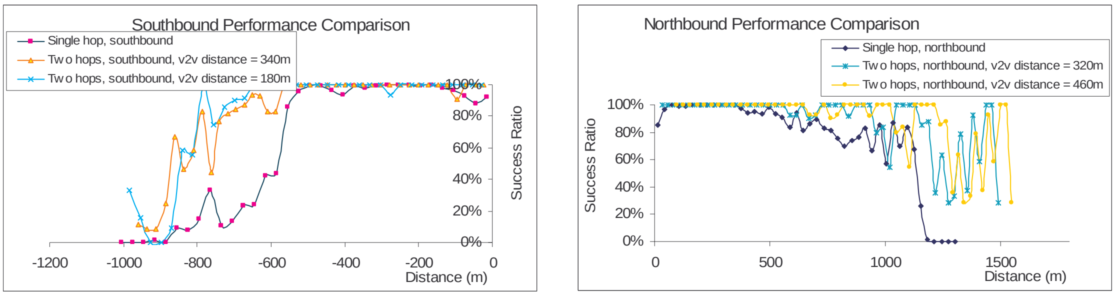
In previous sections, we have observed that road curves and obstructions (e.g., trees and bridges) affect the communication between a moving vehicle and a roadside station. Leveraging V2V communications is a viable approach to mitigate dead spots and extend communication range. For example, a vehicle with poor communication with the roadside station can forward data through another vehicle that has a good connection to the station. We designed multi-hop communication experiments to demonstrate this concept and to measure the degree that the effective communication range could be increased. In these experiments, the receiver was a roadside station placed on the southern side of Peachtree Battle Bridge with its antenna mounted on a tripod. The sender was a moving vehicle looping between Exit 252B and Exit 255 of I-75 and constantly sending data to the receiver. The sender was always behind another vehicle that served as a potential forwarder. Based on the V2V communication results presented in the previous section, we kept the distance between the forwarder and the sender to within 500m, as long as safety conditions could be maintained.

For this set of experiments, we need a forwarding algorithm to achieve the following behavior. Initially, the sender is far from the receiver and the forwarder forwards packets to the receiver for the sender. Once the sender moves into the direct communication range of the receiver, the sender directly communicates with the receiver and the forwarder should stop forwarding packets. Several existing routing algorithms in vehicular networks can be adapted for this purpose [79, 87]. For this study, we have implemented a contention-base geographic forwarding algorithm.

The contention-based scheme works as follows. Every packet is assigned an owner during the process of delivery. The owner of a packet may change over time. The source of a packet initially assumes ownership. The packet owner attaches its current location and broadcasts the packet as an invitation for contention. Vehicles receiving the packet compare their own locations with the location specified in the packet. If their locations are closer to the destination, they join the contention and set a backoff timer. The destination acknowledges the packet immediately. If a contender does not hear the packet (or ack) again before its backoff timer expires, it assumes it wins the contention and becomes the new packet owner and then resends the packet as a contention invitation. Otherwise it drops the contention. The backoff time is computed with $t_{backoff} = (1 - \frac{d}{r}) * t_{max}$ where d is the distance to the sender, r is the radio range and t_{max} is the maximum backoff time. This means that a vehicle closer to the destination sets a shorter backoff timer and is more likely to win the contention. Therefore the ownership is transferred in a way to make the greatest progress.

This algorithm is sufficient for our test case. When the sender is far from the receiver, only the forwarder can receive the packet and will retransmit it after its timer expires. Once the receiver (roadside station) can hear the sender directly, it will acknowledge the packet immediately and thus suppress the retransmission of the forwarder. It is possible that the receiver may receive duplicate packets since the forwarder may not receive the ack. When this happens, the receiver simply drops any duplicate. This algorithm is also applicable if the experiments were to be extended to include a fleet of vehicles.

This forwarding algorithm has been implemented as a Linux kernel module operating in the IP layer so that applications (e.g., IPerf) can run without modification. In our experiments, r is set to 800m and t_{max} to 30ms.



(a) (b)

Figure 48: The comparison of one-hop and two-hop communication performance between a moving vehicle and a roadside station as a function of the distance between the sender and the roadside station

In Figure 48, we plot the one-hop and two-hop communication performance when vehicles travel southbound and northbound, respectively. The one-hop performance is an aggregation of 5 laps. Two laps of two-hop communications are plotted separately each way. Note that the southbound performance is only plotted to the north of Peachtree Battle Bridge (negative distance) while the northbound performance is only plotted to the south of Peachtree Battle Bridge (positive distance). The benefit of extended communication range is obvious: not only can the sender communicate with the receiver in areas out of the range of one-hop communications (e.g., beyond 1200m to the south of the bridge), but also the performance in areas with poor one-hop communications is improved because the receiver can receive from either the sender or the forwarder. Multi-hop forwarding helps offset the impact of the bridge over Exit 254 as well as the road curves to the south of Peachtree Battle Bridge. Through studying receiver traces, we find that the receiver initially only receives from the forwarder, then there is a period when the receptions from both the sender and the forwarder are interleaved, and finally the reception from the sender dominates. This agrees with our expectations.

5.6 Discussion

Through one-hop V2R and V2V experiments, we have shown that good communication performance is possible over a long range (more than 500m) under open road conditions, given our experimental configurations. Previous measurement studies report much shorter communication range. Ott and Kutscher [63] observed a production phase of only about 100m in each side of a roadside base station. We speculate the difference may come from several sources. First, in their experiments, the roadside station was level with the road, while we placed it at an elevated position above the road so that it could (for the most part) maintain a line of sight with vehicles until curves or bridges obstructed its view. Second, they transmitted data at 11Mbps while we transmitted at 2 Mbps, with lower data rate allowing higher power per data symbol and thus longer transmission range. In addition, we used external antennas on both ends of the communication while they only placed the antenna on the moving vehicle only.

Road curves and obstacles degrade communication performance, resulting in much shorter communication range. This suggests that other obstacles, such as large trucks, may result in significantly reduced communication performance. During our experiments, we did not observe that vehicle traffic imposed a significant impact on V2R communications, likely due to the elevated placement of the roadside station, but did affect V2V communications as shown by the significant impact of inter-vehicle distance on communication performance. Further experiments can be designed to explore the vehicle traffic on communication performance. We have shown multi-hop V2V communications can be leveraged to mitigate these adverse factors. Another approach is to deploy relays close to the poor coverage areas, e.g., in the vicinity of roadway curvature. This will be particularly useful where vehicles are expected to be sparse.

The North America DSRC is based on 802.11a which is known to have much shorter range than 802.11b. However, DSRC has been designed with different communication ranges (as long as 1000m) when operating in different power levels. Even though the actual DSRC system may behave differently from the 802.11b devices used here due to different operating frequency and modulation schemes, we still expect our experiments would give some useful indication on how a realistic vehicular environment may affect DSRC performance.

We have seen how communication performance may vary over time and location as vehicles move, which has significant implications on protocol and application design. Any efficient protocol/application must be aware of and adapt to this variation. For example, the video captured by cameras equipped on vehicles can be forwarded backward to improve the vision of drivers. Such a system can adapt to sudden loss of communication by starting collecting information from surrounding vehicles. Fortunately, detailed GIS information concerning the road makes performance variation due to road environment relatively easy to predict (modulo mobile obstacles such as trucks that obstruct LOS) and thus adaptation is easier to implement. Much more difficult is how to adapt to dynamic vehicle traffic. This is an area of future research. It is very likely that future “smart” vehicles will be equipped with multiple radio interfaces (e.g., cellular, 802.11x, DSRC, etc), making it possible to overcome the poor performance of one radio in a specific location [52].

5.7 Conclusion

In this chapter, we have presented our experiences in measuring short-range communication performance between vehicles and between vehicles and roadside stations on a highway. In addition to separation distance, roadway infrastructure interference on V2R performance is identified, e.g., overpasses, curvature (line of sight), trees, etc. V2R performance is predictable and such information can be used in roadside network design. V2V results are dominated by inter-vehicle separation. But, roadway infrastructure interference is also noted. Roadside interference factors can be mitigated by exploiting V2V hops. Similar studies should be conducted for a broad range of settings (e.g., varying packet size, different locations, etc.) to provide a comprehensive evaluation. Another important area of work is to characterize and identify other factors (e.g., weather and vehicle traffic) that may impact communication performance.

CHAPTER

6 CONCLUSION

The vehicular network is an integrated radio network leveraging various wireless technologies that (ideally) work together in a seamless fashion. Several wireless technologies exist for creating vehicular networks, offering different tradeoffs in cost and performance. As a result, different types of vehicular networks may be developed, each exhibiting distinct properties. This thesis identifies various types of vehicular networks, and addresses both their evaluation and design. The major contributions of this thesis are summarized below.

6.1 Vehicular Network Evaluation

As deployment of smart vehicles becomes more widespread, the application of system evaluation methodologies to networked in-vehicle computing systems becomes increasingly more important. This thesis describes our experiences in evaluating vehicular networks using statistical analysis, simulation, and field experiments. The context for these investigations concerns the exploration of the use of vehicle-to-vehicle and vehicle-to-infrastructure communications to disseminate information in vehicular networks.

Various modes of wireless communications may exist in vehicular networks, including WWAN and WLAN using roadside access points for V2R communications, short-range V2V communications, Bluetooth or Ethernet connecting in-vehicle devices, and WiMAX or wired links connecting roadside facilities to the backbone.

Short-range communications (e.g., DSRC) are designed to facilitate V2V and V2R communications and are an important component of ITS systems [83, 84]. Many proposed safety applications [59, 93] rely on short-range communications, e.g., cooperative driving. Other applications might also take advantage of short-range communications, e.g., toll collection, traffic light control, etc. One-hop communication is often sufficient for these applications. Much research has investigated the performance [74] [63] [60] and reliability [80, 93, 95] of short-range communications. In Chapter 5, we presented a measurement study to examine the impact of road environment on short-range communications. The results show that there is no clear cut on whether communication is possible or not as the communication is affected by many factors on the road, including road curves, obstructions, vehicle location, etc. As a result, the communication performance is not uniform on the road. Coupled with the fact that vehicles are moving fast, it can be expected that the communication performance experienced by an individual vehicle may vary significantly over time. However, the communication performance is consistent when one end of communication is fixed, suggesting performance profiles can be developed to optimize communication systems.

Ad-hoc networks utilizing short-range communications can be constructed to disseminate information over a wide area, offering an alternative to the more expensive infrastructure-based services. However, this deployment has many limitations due to uncertainties caused by vehicle movement, insufficient instrumented vehicle density and unreliable wireless channels. Chapter 2 studied spatial propagation of information using multi-hop communications. Analytical models were developed to derive the expected information propagation speed under various vehicle traffic characteristics (e.g., vehicle density and vehicle speed distribution). A simulation-based evaluation was presented to illustrate information propagation along a highway. The results show that information can usually obtain an end-to-end connected path of several miles even when the penetration ratio is as low as 0.2 during the high vehicle traffic periods (e.g., daytime). However, such an E2E path typically does not exist during nighttime even if all vehicles are instrumented. Therefore, multi-hop communications remains a viable and low cost approach for applications that can tolerate some delay or loss (e.g., tourist information), or for disseminating time-sensitive information (e.g., traffic information) during high vehicle traffic periods.

Roadside infrastructures are required in areas where it is desired to provide travelers reliable communication services. WWAN base stations and WLAN access points along the road can provide the required coverage. WWAN last-hop, WLAN last-hop, and multi-hop WLAN are 3 building blocks that can be used to construct vehicular network infrastructure. Chapter 4 identified and analyzed several types of communication architectures (Chapter 4.1) that are suited for providing high bandwidth communications to travelers. These network architectures were also quantitatively assessed under realistic vehicle traffic conditions, through which some insights into designing vehicular network infrastructure were developed.

If continuous connectivity is required, a WLAN last-hop approach might be a good option due to its simplicity, easy deployment and the ability to provide high data throughput. A WWAN last-hop approach is able to provide continuous connectivity but often does not provide sufficient capacity. A hybrid architecture using both WWAN and WLAN is attractive in places where WWAN already exists. It can increase system capacity and is able to provide higher user throughput or support more users while still ensuring ubiquitous connectivity. A major consideration concerns whether multi-hop WLAN should be allowed. As shown in Chapter 4.3, the major tradeoff is whether the reduced cost introduced by multi-hop forwarding can sufficiently justify the additional system complexity. Usually multi-hop forwarding offers limited advantages except in places where there are limitations on the number of access points that can be installed.

If continuous connectivity is not required, a WLAN-based architecture is preferred for quick deployment and relatively low construction cost. The intermittence requirement determines the density of access points that must be deployed.

The design of a vehicular network must address changing vehicle traffic conditions over time. Two major approaches are overprovision and adaptation (Chapter 4.3).

6.2 Design of Effective Data Services

Data services concern the storage, manipulation, aggregation and transport of data inside the network. An effective data service must address the characteristics of the network upon which it will operate on. Chapter 3 discussed the design of enhanced opportunistic forwarding algorithms to address the characteristics of V2V networks. These algorithms rely on in-vehicle processing and inter-vehicle interaction. Time and location based addressing and localized algorithms (in a constantly varying local area) are designed to handle the high dynamics of V2V networks. Data replication is employed to deal with uncertainties inherent in V2V networks. Based on this work, we can derive several design principles of general implication:

Localized processing. Since vehicular networks are highly dynamic and information often has local relevance, large-scale logical structures (e.g., trees) are often impossible, undesirable, or unnecessary to maintain. Rather, localized algorithms that intend to store location-relevant data locally (floating vehicles or roadside stations) are preferred to increase data availability and improve system scalability.

Bring computing close to the data. Compared with sensor networks, vehicular networks represent a resource-rich computing environment. Vehicles will entail more and more advanced sensing capabilities. This means that a large amount of information will be produced by vehicles. It is becoming a challenge as to how to organize this large amount of data to make it more useful and accessible. However, increasingly sophisticated in-vehicle computing systems offer significant data processing capabilities, which can be used to fulfill the vision of "bringing computation close to the data". This offers a way to alleviate this problem.

Deal with uncertainties. Unreliable communication channels, in-vehicle system failures, high vehicle mobility, and network partitioning introduce uncertainties in vehicular networks. Data replication and diversity can be employed to improve performance or increase data availability and reliability. However, too much replication can decrease system efficiency. Thus it is necessary to find the proper balance to keep the system functional.

6.3 Future Direction

Further evaluation should be conducted in a broad range of settings to provide a more comprehensive understanding. Architecture evaluation needs to be performed for vehicle traffic conditions in different time and areas, different data traffic models, communication parameters, etc. It is also necessary to study how network protocols and applications, e.g., TCP, might be affected by proposed network architectures. Our measurement results in Chapter 5 and the previous works [74] [63] [60] gave a general picture of short-range communication performance. Future work could focus on identifying other factors (e.g., weather and vehicle traffic) that can impact communications. It is also important to understand the performance of other communication services (e.g., cellular services) in a vehicular environment. More complex experiments need to be designed to assess proposed services. In particular, experiments can be designed, with the cooperation of government partners, to demonstrate the feasibility of exploiting information and computing power provided by smart vehicles to improve traffic conditions. One example is to exploit the path information provided by vehicles to optimize traffic light timing strategies.

Evaluation methodologies are essential to understand system behaviors as well as to assess alternate approaches toward realizing a rich variety of computing and information services to travelers. Evaluation methodologies are expected to develop in the following areas. In the area of statistical analysis, vehicle traffic flow theory [29] [53] may play a crucial role since it reveals the dynamics of vehicle traffic. Vehicle traffic flow theory can be combined with conventional mathematical tools for studying computer systems (e.g., queuing theory, stochastic process) in deriving useful models. An imminent issue is to determine how much detail about vehicle traffic is needed in analyzing communication systems. For example, vehicle acceleration might not be necessary since information transmission is so fast. In simulations, information needs to flow from the communication simulator to the transportation simulator in order to study the impact of information services on driving behavior and vehicle traffic conditions. While non-safety related services can be modeled using the simulation test bed presented in this thesis, much more detailed simulation, e.g., time scales as low as milliseconds, is required for safety-related applications. Large-scale simulation is necessary to study the system behavior in a wide area. Simulation performance and scalability will then become a crucial factor. Future smart vehicles will be instrumented with more and more devices (e.g., wireless interfaces, cameras, sensors, etc.) These increasingly sophisticated smart vehicles together with roadside infrastructure are required to support advanced applications.

Future research might focus on a framework for efficient data processing following the principles presented in the preceding section. Many challenging research problems will arise. One problem is the delivery of fresh contextual and location-aware information requiring assured quality. For instance, timely and reliable delivery of safety information is critical for drivers. However, drivers should not be distracted by unrelated information. A highly dynamic system consisting of moving vehicles and wireless networks presents many challenges for quality assured treatment, transmission and delivery of information. Another challenging problem is designing services adaptive to varying network resources and vehicle traffic conditions. One solution to this problem is to leverage multiple communication capabilities a vehicle has at its disposal.

As revealed through our research, multiple modes of communications can be exploited to improve the communication performance experienced by a vehicle. A vehicle equipped with multiple radios can be envisioned as a “mobile intranet” on the road. Techniques are required for effective cooperation among multiple radios [11]. In particular, techniques such data striping [7, 36, 53, 67, 75] and vertical handoffs [22, 77] in multi-radio systems as well as mobility management [39] have direct relevance.

On top of network services, a distributed computing middleware can be designed to connect diverse in-vehicle systems manufactured by different automakers and aftermarket suppliers, and roadside resources. This middleware should substantially increase interoperability, encourage resource and data sharing, improve system and data security, ensure the reliability of safety-related applications, and reduce the development and maintenance cost. Grid computing [26] fulfills most of these requirements. However, unlike the conventional "heavyweight" grid system targeting scientific computing applications, this grid platform is required to be extremely lightweight to address the highly dynamic environment.

REFERENCES

1. Developing an Information Technology-Oriented Basic Research Program for Surface Transportation Systems, National Science Foundation and US Department of Transportation, Chicago, IL, 2000.

2. http://www.wi-fiplanet.com/tutorials/article.php/1453261, 2005.

3. http://www.wimaxforum.org/technology/tech_overview, 2005.

4. Not In The Public Interest - The Myth of Municipal Wi-Fi, The New Millennium Research Council Report, 2005.

5. Report of NSF Workshop on Fundamental Research in Networking, 2003.

6. *Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle Systems - 5GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 2003.

7. Adiseshu, H., Parulkar, G. and Varghese, G., A Reliable and Scalable Striping Protocol. in *SIGCOMM'96*, (1996).

8. Aguayo, D., Bricket, J., Biswas, S., Judd, G. and Morris, R., Link-level Measurements from an 802.11b Mesh Network. in *SIGCOMM*, (2004).

9. ALANR. Iperf, 2004.

10. Anastasi, G., IEEE 802.11 Ad Hoc Networks: Performance Measurements. in *ICDCSW*, (2003).

11. Bahl, P., Adya, A., Padhye, J. and Wolman, A. Reconsidering Wireless Systems with Multiple Radios. *SIGCOMM Computer Communication Review (CCR)*.

12. Bahl, P., Chandra, R. and Dunagan, J., SSCH: Slotted Seeded Channel Hopping for Capacity Improvement in IEEE 802.11 Ad-Hoc Wireless Networks. in *MobiCom'04*, (2004).

13. Bai, F., Sadagopan, N. and Helmy, A. The IMPORTANT framework for analyzing the Impact of Mobility on Performance Of RouTing protocols for Adhoc NeTworks. *Elsevier Ad Hoc Networks*.

14. Broch, J., Maltz, D.A., Johnson, D.B., Hu, Y.-C. and Jetcheva, J., A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols. in *MobiCom*, (1998).

15. California PATH project. http://www.path.berkeley.edu/, 2005.

16. Chang, X., Network simulations with OPNET. in *Winter Simulation Conference*, (1999).

17. Chen, Z.D., Kung, H. and Vlah, D., Ad Hoc Relay Wireless Networks over Moving Vehicles on Highways. in *MobiHoc*, (2001).

18. Cheng, Y.-c. and Robertazzi, T.G. Critical Connectivity Phenomena in Multihop Radio Models. *IEEE Trans. on Communications*, *37*. 770-777.

19. Davis, J., Fagg, A. and Levine, B., Wearable computers as packet transport mechanisms in highly-partitioned ad-hoc networks. in *International Symposium on Wearable Computing*, (2001).

20. Dogan, A., Korkmaz, G., Liu, Y., Ozguner, F., Ozguner, U., Redmill, K., Takeshita, O. and Tokuda, K., Evaluation on Intersection Collision Warning System Using Inter-vehicle Communication Simulator. in *7th International IEEE Conference on Intelligent Transportation systems*, (2004).

21. Dousse, O., Thiran, P. and Hasler, M., Connectivity in ad-hoc and hybrid networks. in *Infocom 2002*, (2002).

22. Dutta, A., Kim, B., Zhang, T., Baba, S., Taniuchi, k., Ohba, Y. and Schulzrinne, H., Experimental Analysis of Multi Interface Mobility Management with SIP and MIP. in *IEEE Wirelesscom*, (2005).

23. Fall, K., A Delay-Tolerant Netowork Architecture for Challenged Internets. in *SIGCOMM'03*, (2003).

24. Federal Highway Administration. Traffic Software Integrated System User's Guide, 2003.

25. FleetNet. www.et2.tu-harburg.de/fleetnet, 2004.

26.Foster, I. and Kesselman, C. *The Grid 2: Blueprint for a New Computing Infrastructure*, 2004.

27. Fübler, H., Widmer, J., Kasemann, M., Mauve, M. and Hartenstein, H. Contention-Based Forwarding for Mobile Ad-Hoc Networks. *Elsevier's Ad-hoc Networks*, *1* (4). 351-369.

28. Fujimoto, R., Wu, H., Guensler, R. and Hunter, M., Evaluating Vehicular Networks: Analysis, Simulation, and Field Experiments. in *Symposium on Modeling and Simulation Tool for Emerging Telecommunications Networks: Needs, Trends, Challenges, Solutions*, (Munich, Germany, 2005).

29. Gerlough, D.L. and Huber, M.J. *Traffic flow theory: a monograph*. Washington: Transportation Research Board, National Research Council, 1975.

30. Grossglauser, M. and Tse, D., Mobility increases the capacity of ad-hoc wireless network. in *Infocom 2001*, (2001).

31. Gupta, P. and Kumar, P.R. The Capacity of Wireless Networks. *IEEE Transactions on Information Theory*, *46* (2). 388-404.

32. Heinzelman, W., Kulik, J. and Balakrishnan, H., Adaptive Protocols for Information Dissemination in Wireless Sensor Networks. in *ACM Mobicom'99*, (1999).

33. Hsieh, H.-Y. and Sivakumar, R., Internetworking WWANs and WLANs in next generation wireless data networks. in *3G Wireless and Beyond*, (2002).

34. Hsieh, H.-Y. and Sivakumar, R., On Using the Ad-hoc Network Model in Cellular Packet Data Networks. in *MOBIHOC'02*, (2002).

35. Hsieh, H.-Y. and Sivakumar, R., Performance comparison of cellular and multi-hop wireless networks: A quantitative study. in *SIGMETRICS'01*, (2001).

36. Hsieh, H.-Y. and Sivakumar, R., A Transport Layer Approach for Achieving Aggregate Bandwidths On Multi-homed Mobile Hosts. in *Mobicom*, (2002).

37. Intanagonwiwat, C., Govindan, R. and Estrin, D., Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks. in *ACM MobilCom'00*, (2000).

38. Johnson, D.B. and Maltz, D.A. Dynamic source routing in ad hoc wireless networks. *Mobile Computing*.

39. Jonsson, U., Alriksson, F., Larsson, T., Johansson, P. and Jr., G.Q.M., MIPMANET-Mobile IP for Mobile Ad Hoc Networks. in *MOBIHOC*, (2001).

40. Juang, P., Oki, H., Wang, Y., Martonosi, M., Peh, L.-S. and Rubenstein, D., Energy-Efficient Computing for Wildlife Tracking: Design Tradeoffs and Early Experiences with ZebraNet. in *ASPLOS*, (2002).

41. Karp, B. and Kung, H.T., GPSR: Greedy Perimeter Stateless Routing for Wireless Networks. in *MobiCom*, (2000).

42. Khelil, A., Becker, C., Tian, J. and Rothermel, K., An Epidemic Model for Information Diffusion in MANETs. in *MSWiM'02*, (2002).

43. Kutzner, K., Tchouto, J.-J., Bechler, M., Wolf, L., Bochow, B. and Luckenbach, T., Connecting Vehicle Scatternets by Internet-Connected Gateways. in *MMC'2003*, (2003).

44. LANL. Transportation Analysis Simulation System (TRANSIMS), 1999.

45. Lee, J., Hunter, M., Ko, J., Guensler, R. and Kim, H.K., Large-Scale Microscopic Simulation Model Development Utilizing Macroscopic Travel Demand Model Data. in *6th Transportation Speciality Conference*, (Toronto, Ontario, Canada, 2005).

46. Lettieri, P. and Srivastava, M.B., Adaptive frame length control for improving wireless link throughput, range, and energy efficiency. in *INFOCOM*, (1998).

47. Li, J., Blake, C., Couto, D.S.J.D., Lee, H.I. and Morris, R., Capacity of Ad Hoc Wireless Networks. in *ACM Mobicom*, (2001).

48. Li, Q. and Rus, D., Sending Messages to Mobile Users in Disconnected Ad-hoc Wireless Networks. in *Mobicom'2000*, (2000).

49. Lin, Y.-D. and Hsu, Y.-C., Multihop Cellular: A New Architecture for Wireless Communications. in *IEEE INFOCOM 2000*, (2000).

50. Lochert, C., Hartenstein, H., Tian, J., Fubler, H., Hermann, B. and Mauve, M., A Routing Strategy for Vehicular Ad Hoc Networks in City Environment. in *IEEE IV'2003*, (2003).

51. Lott, M., Performance of a Medium Access Scheme for Inter-vehicle. in *SPECTS'02*, (2002).

52. Luo, H., Ramjee, R., Sinha, P., Li, L. and Lu, S., UCAN: A Unified Cellular and Ad-Hoc Network Architecture. in *MobiCom'03*, (2003).

53. Magalhaes, L. and Kravets, R., Transport Level Mechanisms for Bandwidth Aggregation on Mobile Hosts. in *ICNP'01*, (2001).

54. Mauve, M., Widmer, J. and Hartenstein, H. A Survey on Position-Based Routing in Mobile Ad Hoc Networks. *IEEE Network*, *15* (6).

55. May, A.D. *Traffic Flow Fundamentals*. Prentice-Hall Inc., Englewood Cliffs, NJ 07632, 1990.

56. McCanne, S. and Floyd, S. The {LBNL} Network Simulator, 1997.

57. McLean, T., Fujimoto, R. and Fitzgibbons, B. Next Generation Real-Time RTI Software *Workshop on Distributed Simulations and Real-Time Applications*, 2001, 4-11.

58. Meguerdichian, S., Slijepcevic, S., Karayan, V. and Potkonhak, M., Localized Algorithms in Wireless Ad-Hoc Networks: Location Discovery and Sensor Exposure. in *MobiHOC 2001*, (2001).

59. Morsink, P., Cseh, C., Gietelink, O. and Miglietta, M., Design of an application for communication-based longitudinal control in the CarTALK project. in *IT Solutions for Safety and Security in Intelligent Transport (e-Safety),* (2002).

60. Moske, M., Fubler, H., Hartenstein, H. and Franz, W., Performance Measurements of a Vehicular Ad Hoc Network. in *VTC/Spring*, (2004).

61. Ni, S.-Y., Tseng, Y.-C., Chen, Y.-S. and Sheu, J.-P., The Broadcast Storm Problem in a Mobile Ad Hoc Networks. in *IMobicom '99*, (1999).

62. Niculescu, D. and Nath, B., Trajectory Based Forwarding and Its Applications. in *MobiCom'03*, (2003).

63. Ott, J. and Kutscher, D., Drive-thru Internet: IEEE 802.11b for "Automobile" Users. in *INFOCOM'04*, (2004).

64. Papadopouli, M. and Schulzrinne, H., Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices. in *ACM MobiHoc*, (2001).

65. Pentland, A., Fletcher, R. and Hasson, A. DakNet: Rethinking Connectivity in Developing Nations. *IEEE Computer*.

66. Qiao, C., Wu, H. and Tonguz, O., iCAR: An integrated cellular and ad-hoc relay system. in *IEEE International Conference Computer Communication Network*, (2000).

67. Qureshi, A. and Guttag, J., Horde: Seperating Network Striping Policy from Mechanism. in *Mobisys,* (2005).

68. Ritter, M.W. The Future of WLAN *ACM Queue*, 2003.

69. Ross, S.M. *Introduction to Probability Models*. Academic Press, 2000.

70. Rudack, M., Meincke, M. and Lott, M., On the Dynamics of Ad Hoc Networks for Inter Vehicle Communications (IVC). in *ICWN'02*, (2002).

71. Russo, K.L., Shuette, L.C., Smith, J.E. and McGuire, M.E. Effectiveness of Various New Bandwidth Reduction Techniques in ModSAF. in *Proceedings of the 13th Workshop on Standards for the Interoperability of Distributed Simulations*, 1995, 587-591.

72. Scalable Network Technologies. QualNet, 2004.

73. Shah, R.C., Roy, S., Jain, S. and Brunette, W., Data MULES: Modeling a Three-tier Architecture for Sparse Sensor Networks. in *SNPA*, (2003).

74. Singh, J.P., Bambos, N., Srinivasan, B. and Clawin, D., Wireless LAN Performance under varied stress conditions in vehicular traffic scenarios. in *IEEE VTC Fall 2002*, (2002).

75. Sivakumar, H., Bailey, S. and Grossman, R.L., PSockets: The Case for Application-level Network Striping for Data Intensive Applications using High Speed Wide Area Networks. in *IEEE Super Computing*, (2000).

76. Sivakumar, R., Sinha, P. and Bharghavan, V. CEDAR: a Core-Extraction Distributed Ad hoc Routing algorithm. *IEEE Journal on Selected Areas in Communication*, *17* (8).

77. Stemm, M. and Katz, R.H. Vertical Handoffs in wireless overlay networks. *ACM Mobile Networks and Applications*, *3* (4).

78. Tian, J. and Rothermel, K. Building Large Peer-to-Peer Systems in Highly Mobile Ad Hoc Networks: New Challenges? *Technical Report 2002/05*, University of Stuttgart, 2002.

79. Tian, J., Stepanov, I. and Rothermel, K. Spatial Aware Geographic Forwarding for Mobile Ad Hoc Networks, University of Stuttgart, 2002.

80. Torrent-Moreno, M., Jiang, D. and Hartenstein, H., Broadcase Reception Rates and Effects of Priority Access in 802.11-Based Vehicular Ad-Hoc Networks. in *VANET 2004*, (2004).

81. Vahdat, A. and Becker, D. Epidemic routing for partially-connected ad hoc networks, Duke University, 2000.

82. Van Aerde, M. and Yager, S. Dynamic Integrated Freeway/Traffic Signal Networks: Problems and proposed Solutions. *Transportation Research A*, *22* (6). 435-443.

83. Werner, J. More Details Emerge about the VII Efforts, Newsletter of the ITS Cooperative Deployment Network, http://www.ntoctalks.com/icdn/vii_details_itsa04.html, 2004.

84. Werner, J. USDOT Outlines the New VII Initiative at the 2004 TRB annual meeting, Newsletter of the ITS Cooperative Deployment Network, http://www.nawigts.com/icdn/vii_trb04.html, 2004.

85. Wischhof, L., Ebner, A., Rohling, H., Lott, M. and Hafmann, R., Adaptive Broadcast for Travel and Traffic Information Distribution Based on Inter-Vehicle Communication. in *IEEE IV'2003*, (2003).

86. Wolfram Research. Mathematica, 2003.

87. Wu, H., Fujimoto, R., Guensler, R. and Hunter, M., MDDV: A Mobility-Centric Data Dissemination Algorithm for Vehicular Networks. in *1st ACM workshop on vehicular ad hoc networks*, (2004).

88. Wu, H., Fujimoto, R., Hunter, M. and Guensler, R., An Architecture Study of Infrastructure-base Vehicular Networks. in *MSWiM'05*, (2005).

89. Wu, H., Fujimoto, R. and Riley, G., Analytical Models for Information Propagation in Vehicle-to-Vehicle Networks. in *IEEE VTC 2004-Fall*, (2004).

90. Wu, H., Lee, J., Hunter, M., Fujimoto, R., Guensler, R. and Ko, J., Simulated Vehicle-to-Vehicle Message Propagation Efficiency on Atlanta's I-75 Corridor. *Transportation Research Record (TRR)*, (2005).

91. Wu, H., Palekar, M., Fujimoto, R., Guensler, R., Hunter, M., Lee, J. and Ko, J., An Empirical Study of Short Range Communications for Vehicles. in *2nd ACM workshop on vehicular ad hoc networks*, (2005).

92. Wu, H., Qiao, C., De, S. and Tonguz, O. Integrated Cellular and Ad Hoc Relaying Systems. *IEEE Journal on Selected Areas in Communication*, *19* (10).

93. Xu, Q., Mak, T., J.Ko and Sengupta, R., Vehicle-to-Vehicle Safety Messaging in DSRC. in *ACM VANET 2004*, (2004).

94. Xu, Q., Sengupta, R. and Jiang, D., Design and Analysis of Highway Safety Communication protocol in 5.9 GHz Dedicated Short Range Communication Spectrum. in *IEEE VTC'03*, (2003).

95. Yin, J., ElBatt, T., Yeung, G. and Ryu, B., Performance Evaluation of Safety Applications over DSRC Vehicular Ad Hoc Networks. in *VANET 2004*, (2004).

96. Yoon, J., Liu, M. and Noble, B., Sound Mobility Model. in *MobiCom'03*, (2003).

97. Zhao, W. and Ammar, M., Message Ferrying: Proactive Routing in Highly-Partitioned Wireless Ad Hoc Networks. in *9th IEEE Workshop on Future Trends in Distributed Computing Systems*, (2003).

98. Zhao, W., Ammar, M. and Zegura, E., A Message Ferrying Approach for Data Delivery In Sparse Mobile Ad Hoc Networks. in *MobiHoc'04*, (2004).

99. Ziliaskopoulos, A.K. and Zhang, J., A Zero Public Infrastructure Vehicle Based Traffic Information System. in *TRB 2003 Annual Meeting*, (2003).